

INSTITUTTET  
FOR SAMMENLIGNENDE  
KULTURFORSKNING



SERIE A: FORELESNINGER

VIII

BERNHARD KARLGREN:  
PHILOLOGY AND ANCIENT CHINA

OSLO 1926

H. ASCHEHOUG & CO. (W. NYGAARD)

LEIPZIG

OTTO HARRASSOWITZ

PARIS

HONORÉ CHAMPION

LONDON

WILLIAMS & NORGATE, LTD.

CAMBRIDGE, MASS.

HARVARD UNIVERSITY PRESS

INSTITUTTET  
FOR SAMMENLIGNENDE KULTURFORSKNING

---

# PHILOLOGY AND ANCIENT CHINA

BY  
BERNHARD KARLGREN



495-114

Kar

27454

OSLO 1926

H. ASCHEHOUG & CO. (W. NYGAARD)

LEIPZIG

OTTO HARRASSOWITZ

PARIS

HONORÉ CHAMPION

LONDON

WILLIAMS & NORGATE, LTD.

CAMBRIDGE, MASS.  
HARVARD UNIVERSITY PRESS



CENTRAL ARCHAEOLOGICAL  
LIBRARY, NEW DELHI.

Acc. No. .... 495.114 ..... 1 Kar

Date.....

Call No. .... 495.114 / Kar

PRINTED IN SWEDEN

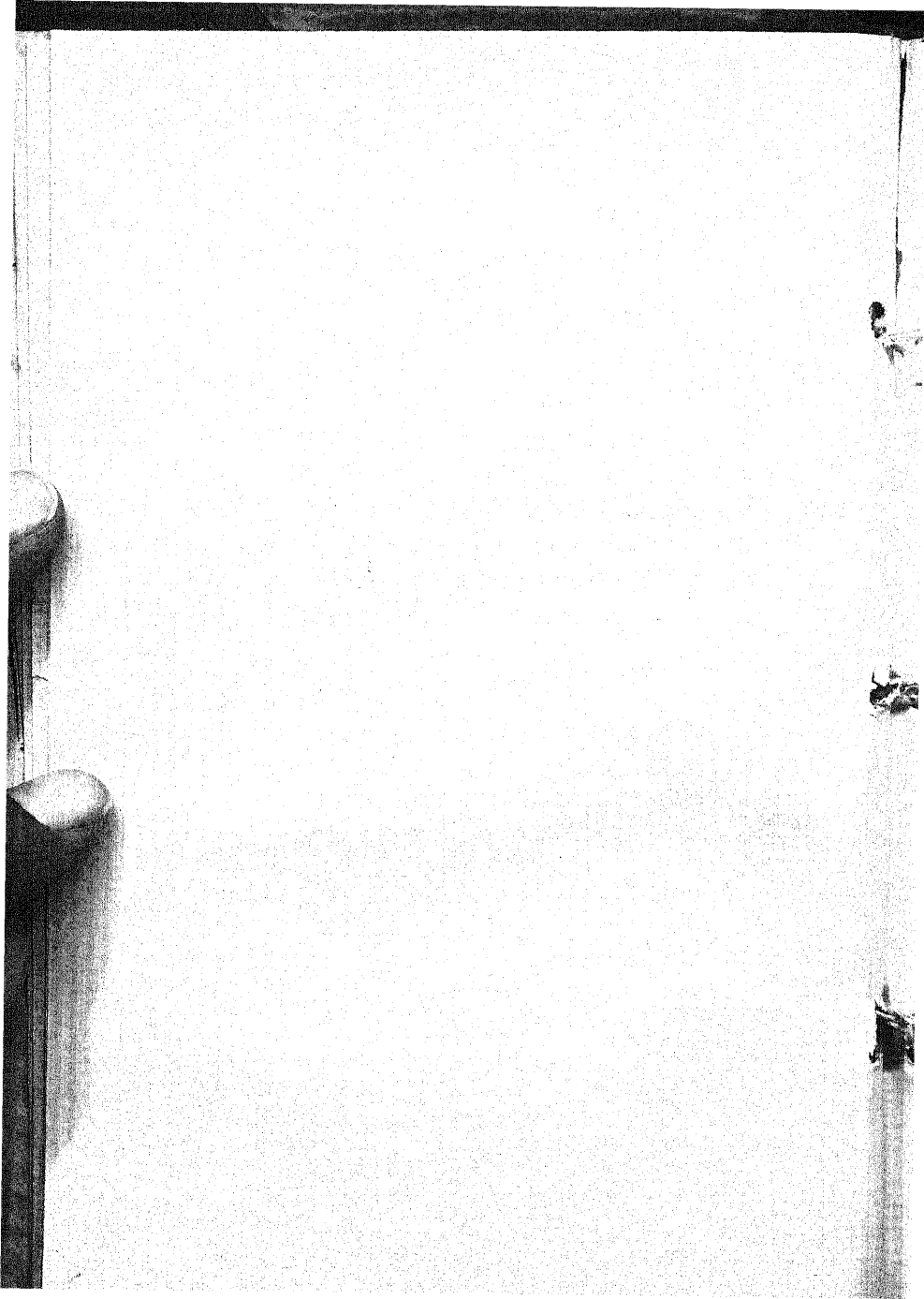
GÖTEBORG 1926  
ELANDERS BOKTRYCKERI AKTIEBOLAG

To

*J. A. Lundell*

in gratitude and friendship.

B. K.



## INTRODUCTION.

When the learned world in the capital of Norway set itself the task of striking a great blow for the humanities, and solved the problem by founding an Institute for Comparative Research in Human Culture, this was in a preeminent degree in accordance with the scientific spirit of the times. Originally comprising very narrow objects of research, classical antiquity and the national branches of knowledge, the humanities first began to make good progress when the field of view was widened, when new domains of research were opened, when the thought and language world of India was brought into the focus of investigation and placed in relationship to the branches of knowledge that had previously been studied. And this striving to widen the area of experience, and then throw light on it by comparative methods, runs like a red filament through all modern humanistic studies, whether in the historical or in the philological schools. So, too, research has been extended to the old civilizations of America, and to its ethnology and present-day languages, on the one hand, and to the motley civilizations of Asia in old and modern times, with its great diversity of languages, on the

other. For the historian of the North, for instance, the Japanese Viking Age offers attractive parallels, and our great Swedish historian, Harald Hjärne, has delved with erudition and acumen in the histories of both Japan and Sweden. For the philosopher Chinese thought, even if it cannot measure itself with the cream of Indian thought, should offer much of interest. And for the philologist the East has vast linguistic fields from which rich veins may be mined, which, once they have been worked and the results re-examined, will assuredly play a great part in helping to fix the general laws that underlie the evolution of languages.

This striving for expansion and for comparative studies, which is the fundamental idea that has called into being this Norwegian Institute, has created a desire, even at this early stage of its existence, to see represented here a branch of research, the East Asiatic, which is comparatively young and which is still almost a stranger in our northern universities. It is with pleasure and gratitude that I have accepted your invitation that summons me to tell you something about the work done in this field of research.

A lecturer at this Institute should never lose sight of the conception, comparative studies. It should therefore be my task to describe some important aspect of Sinology, so that the parallels and contrasts to phenomena in other branches of research may be revealed as clearly as possible. My own studies have been devoted hitherto chiefly to the history of Chinese, especially its phonology and the evolution of its sounds, and this might have tempted me to choose as the subject of my lectures a detailed description of this, in reality very

fertile field. It would afford linguists<sup>1</sup> a copious gleanings which would reveal exactly the same phenomena as those we have in the Indo-European languages — palatalising and velarising under certain conditions, assimilation and dissimilation, the influences of analogy, the dropping of consonants, phenomena connected with the expiratory and musical accents — in short, parallels that will be of the greatest value when we approach the task of trying to formulate comprehensive evolutive phonetics. Besides this there would, of course, be various things of a special nature, peculiar to the development of the Chinese language, and deviations from prevalent phenomena, which might be very instructive.

But I am sure that such an exposition, specially bearing on pure linguistics, would not fulfil the purpose of the series in a satisfactory way, i. e. to give the humanist in general an idea of the analogies or differences that Sinology presents, compared with other subjects of research; I must make my exposition less specialized. Nevertheless, matters are so fortunately situated that I need not doff the linguistic gown, and don that of the historian or of the historian of literature to attain this more general aim. For in China it is a fact that the very nature of the language, its peculiarity, and the original char-

<sup>1</sup>) I wish to state here expressly, that in the following pages I use — in accordance with continental usage — the terms *linguistics*, *linguistic*, *linguist* with regard to the study of language as such (phonology, morphology, syntax, semasiology, studied descriptively, historically, comparatively), and the terms *philology*, *philological*, *philologist* in a wider sense, language study generally, including also language as a conveyance for literature (text critics, exegetics, «realia» etc.). In English there is no consistent terminology in this respect.

acter of the writing conditioned by it, explain much that is strange in the old civilization of China, and even in the new, which is so firmly rooted in the past. It is precisely for linguistic reasons, owing to the unusual nature of the language, that the investigation of Chinese antiquity is in many respects entirely different from corresponding tasks for students of an occidental branch of knowledge. The parallels are many and interesting, but the contrasts are also numerous; to some extent we shall have to adopt quite other methods; and these very divergencies, occasioned by the peculiarity of the language, are instructive and deserve to be set forth in some detail. I therefore take this as my subject:

What constitutes the peculiar nature of the old language of China, and with that, of its literature? What are the philological problems and methods that must be faced to throw light upon them? And what position does the philologist take up towards the effects of these peculiarities in our day, the modern problems of reform in the domain of language and script?

My exposition thus falls into three parts. I will first treat quite summarily the character of the language and script of China — a general description, therefore, of matters that have been known for a long time and have been treated by various writers<sup>1</sup>, but to which I will add some personal points of view, — a description that is indispensable to enable us later on to move more freely. Then I will try to give an idea of the tasks and methods of sinological linguistics, in which connection I will offer an easily grasped summary of my own earlier investigations, as well as some indica-

<sup>1</sup>) E. g., B. Karlgren, *Sound and Symbol in Chinese*. Oxford, 1923.

tions of the lines on which, in my opinion, research should continue. And finally I will enter upon a purely practical sphere, and briefly touch on the problem which at present is greatly troubling the leading intellects, not only in China but also in its cultural colony, Japan; the problem of how to get free in the modern tongue and script from the China of the past, which still remains like a skeleton in the cupboard.



## CHAPTER I.

Let us begin by casting a glance at a sentence in modern Mandarin, the dialect that is spoken in northern and central China, and in the variety of this dialect spoken in the capital, Peking.

*t'a chē-ko p'êng-yu shǐ ko hên ts'ung-ming*  
He this piece friend-friend be piece very hearing-seeing  
*jên, k'ü nien tao-liao si-yang ye hüe shuo*  
man go year reach-complete west ocean also learn speak  
*ko kuo hua*  
each country language.

«This friend of his is a very clever man; last year, when he went to Europe, he learnt to speak the languages of several countries.»

If we examine this specimen in the first place from the point of view of phonology, we at once notice a great simplicity, — always excepting the tones, the musical accent, which I have not marked, and which somewhat complicates matters. A striking feature is that Pekinese does not suffer consonant groups, or several consonants together (affricates like *ts-*, *ch-*, *ch'-*, are really single sounds, which is why they are often written in the West with a single letter, as in German *Zahl*, etc.). The scheme for one syllable is:

*a* or *xa* or *ax* or *xax*.

Furthermore it is to be noted that no words can begin with voiced plosives, *b-*, *d-*, *g-*, or voiced affricates, as *dz-*, *dj-*. Lastly, no syllable can end in anything but a vowel, *-n*, or *-ng*; hence, *pa*, *pan*, *pang* are possible, whereas *pab*, *pam*, *pas*, *pat*, *pad*, *par* *pal*, *paf*, etc. are quite impossible. Furthermore, we see that the simple words are all of one syllable (for which reason Chinese has been called a monosyllabic language), which accentuates the simplicity and monotony of its linguistic garb. These strong restrictions naturally cause a great lack of range; if the words must be monosyllabic, if several consonants cannot come together, if words may not begin or end in any given consonant — well, it becomes impossible to produce as many differently sounded syllables as there are independent single words in the language. As a matter of fact, Pekinese has only about 420 different syllables, though the number is increased to about 1200 by the various tones, and when all the words of a language must be forced into these few syllables, the result is that many words have the same sound, just as, for instance, in the English *bear* (*nascor*) and *bear* (*ursus*). How they manage to get out of the practical difficulties of this, is a question to which I shall return in another connection; at present we are dealing only with the phonological resources of the language.

So much for the system of sounds. The rigid monotony of the language is, however, still further accentuated by another peculiarity, also to be seen in our specimen. We can observe that when words are to form sentences, they are strung together without their mutual relationship being marked by any inflections. If we take the Latin word *liber* with its forms *librum*, *libri*, *libro*, *libros*, *librorum*,

*libris*, we shall find that all these are represented in Chinese by the single form *shu*. There are, it is true, a few auxiliary words, e. g. a genitive particle *-ti*, so that 'his father' can be rendered by *t'a-ti tie*. But these auxiliary words do not by any means correspond to our regular terminations, for they are not obligatory. In English you cannot at pleasure say *he father* instead of *his father*; but in Chinese it is perfectly satisfactory to say quite simply *t'a tie*, the genitival sense being gathered from the context. The small number of elucidative grammatical auxiliary words, which can be used at discretion, do not therefore invalidate the general conclusion that Chinese is a so-called *isolating* language, i. e. a language that is made up of invariable words which always have the same appearance, whether in combination or isolated. And just as there are no inflectional endings, so also there are no derivative prefixes or suffixes, such as we find in *be-loved*, *for-swear*, *sight-ly*, *sight-li-ness*, *tru-th*, *not-able*. *Hing* means *luck* or *lucky* or *luckily*. *Sin* means *true* or *truth* or *trust*. In a word, modern Pekinese has no morphology.

However, China is a vast country, with a population of between three and four hundred million souls, and it stands to reason that it has innumerable dialects. These are often so dissimilar that they are like foreign tongues to each other; thus, on the south coast of China there is a whole row of dialects, the speakers of which simply cannot understand each other or a north Chinaman. In our little preparatory descriptive account we must therefore not be rash and draw our conclusions solely from the dialect of the capital. Let us take a couple of specimens

from elsewhere. Thus, a sentence in Cantonese runs as follows:

*mui kuok ing hok-fān pīt-kuok-ke hou-ch'ü,*  
each country ought learn-back other country's good-place  
*kām-k'ap yu kái tsī-kei kuo.*  
join-come must avoid self-self fault.

«Each country ought to copy the good points of other countries, and avoid the faults of its own.»

The sounds, as we can see, differ somewhat from those in Pekinese. The dialect of Canton allows *-m*, *-p*, *-t*, *-k* at the end of a word, which is not possible in Pekinese. But like the latter, Cantonese does not allow *b-*, *d-*, *g-*, etc., and also has a very limited number of different syllables — about 720 (always apart from the tones). As in Pekinese, too, consonant groups must not occur, it is monosyllabic, and has also the very same isolating character, invariable forms put together with a few auxiliary words.

If we take the dialect of Shanghai, we shall find a sentence of this kind:

*bīng mā dzung, ye-dze i-sēng m -me*  
illness very heavy actual-be medicine-scholar not not-have  
*māng-deu*  
hope-head

«The illness is very severe; now the doctor has no hope».

Here, it is true, *b-*, *d-*, *g-*, *dj-*, *dz-* are allowed at the beginning of a word, contrary to what was the case in Pekinese and Cantonese, but as in Pekinese *-m*, *-p*, *-t*, *-k* are not tolerated at the end of a word, and monosyllabism and the absence of inflections are the same.

As a matter of fact, the result remains identical, to whichever of the many dialects of China we turn. The phonological resources vary not a little from dialect to

dialect, but the important and decisive features are common to all of them — the absence of consonant groups, final consonants restricted to certain sounds which are few in number, poverty in the number of dissimilar sounding syllables as a result of these conditions (in no dialect above 1000, leaving the tones out of consideration), monosyllabism, and the absence of inflections and derivatives. The facts here referred to give modern Chinese a stamp of excessive simplicity, one is tempted to say primitiveness. It is therefore not at all surprising that in the nineteenth century, when attention was directed for the first time to linguistic families and their characteristics, Chinese was taken as the type of the primitive, undeveloped languages, — those which had not yet attained to the same wealth of inflections, derivatives, and polysyllabic simple words as the European languages.

It was not long, however, before the opposite view began to assert itself, viz., that in an earlier stage Chinese had perhaps had a richer phonology, perhaps consonant-groups, had even possibly been polysyllabic, and had perchance had inflections and derivatives which in the course of time had been ground down and lost. In a paper on the linguistic position of Chinese that appeared in 1881, Wilhelm Grube, a German scholar, became the spokesman of this idea.

The uninitiated will be tempted to put this question: China, as we know, has a very old literature, with some documents that date back to 2,000 years B. C., and has preserved a rich literature from the first millennium B. C.; why then need we speculate about these things? Why not go to the old texts, and see whether the phonology was richer, and whether, in contradistinction to what is the

case now, there were polysyllabic words, inflections, and derivatives in pre-Christian times? It is unfortunately impossible to answer the question in this simple way. For, as we shall soon see in greater detail, the Chinese script is not one of letters or phonetic symbols, but one of ideas, which does not permit us to read off the old texts as they sounded at the time when they were written. Every word has one undivided symbol, a conventional symbol for the meaning of a word, an idea, no matter what fashion of sound this idea had in old times, or has in the modern dialects. As I have said, we will return later to this peculiar script; the chief point for the moment is that the old texts do not furnish us without further ado with the key to the nature of the old Chinese language. Conclusions had to be sought, to begin with, by finding analogies with other languages.

And, in point of fact, we have not far to go to find striking examples of such docking and simplification. In French we can see how a whole number of once dissimilar terminations have been levelled under one sound, *-o*. Thus, *beau*, *chevaux*, *dos*, *trop*, *mot*, *mots*, *haut*, *hauts*, *échafaud*, etc. Now, since Pekinese words can only end in a vowel, or in *-n* and *-ng*, we have every right to suspect that in many of the innumerable words now ending in a vowel, Ancient Chinese had final consonants which have been dropped as in the French words just mentioned; i. e. that a Pekinese *pi* may derive from an older *pit* or *pik* or the like. Just as in the English word *knock*, or in the Swedish *hvit*, the initial consonants have been dropped in pronunciation, we could imagine Ancient Chinese consonant-groups to have become simplified until there are now only simple consonant-sounds left. And, just as in French the forms *je parl(e)*, *tu*

*parl(es)*, *ils parl(ent)* originally had quite dissimilar endings which have now been dropped in the spoken language until only the uninflected stem remains, *parl*; so, too, Chinese may have had suffixes that have been dropped and have disappeared. This is the normal process of development in our Indo-European languages. While the older of them, Greek, Latin and Sanskrit, have a very copious accidence, with a multitude of inflectional elements, we have gradually lost them, so that, for instance, English has become almost as uninflected as Chinese. To a great extent the analytic English language even rejects derivatives. From the adjective *clean* there is the derived verb *to cleanse*, but nevertheless the commonest form of the verb is *to clean*, i.e. the adjectival stem without a suffix. This is precisely as in modern Chinese. Finally, as regards monosyllabism, there are Swedish dialects in which *bonden* has become *bonn*, *slitit* > *slēt*, etc., and there is nothing unreasonable in assuming that, at a very early stage, Chinese also had dissyllabic simple words which were subsequently contracted by syncope into monosyllabic ones.

So much for these calculations of probabilities, which make it a plausible assumption that modern, sound-prescribed, monosyllabic, uninflected Chinese, far from being a primitive, undeveloped language, represents a very advanced, extremely reduced stage, which has gone further in the track of simplification and levelling than even English.

But we cannot, of course, content ourselves with these general arguments, and must attain to concrete results. Every linguist knows that we must adopt the methods of comparative philology in order to make headway. Later on I will show, in some detail, how in spite of the obstacles the Chinese script puts in our path, we can start from

purely Chinese material and get some distance in our investigation of how Ancient Chinese sounded. But at the outset there is no harm in directing our gaze to a more distant horizon, view the other languages with which Chinese may be assumed to be related, and thus get some fixed point, so to speak, at longer range.

The view has prevailed for some considerable time that Chinese and Siamese are sister languages, which together with a number of less important ones in Farther India form one great branch of the Indo-Chinese, or, as I would prefer to call them, the Sinitic languages, the other main branch of which consists of Tibetan, Burmese, and other languages.

Though the relationship between Chinese and Siamese is a distant one, and the word correspondences as at present proved, are still few, interesting parallels are easily discoverable. Siamese is a monosyllabic and «isolating» language like Chinese, but its phonology has one or two remarkable features. Consonant groups, which do not exist in modern Chinese, occur there in profusion, and we find, for instance, that the Cantonese *lām*, meaning 'indigo, blue', is identical with Siamese *k<sup>r</sup>rām*, which derives from an earlier form *grām*. Here we have at once a proof that an older stage of the language had a *gr-* or *gl-* which has been simplified in Chinese to *l-*, just as in English *knock* has become *nock*.

Of far more importance, however, is Tibetan. Very fortunately we have Tibetan phonetically written in the seventh century A. D., in a script based on the Sanskrit alphabet. So we can form a very exact idea of how Tibetan was pronounced 1300 years ago, and have a good and old comparative landmark — always provided that



Tibetan and Chinese are really related. Some scholars are still doubtful on this point (see Henri Maspero in the Bull. Ec. Franç. d'Extr. Orient, 1920, P. 22, note), but it seems to me that the established correspondences are sufficient to prove the relationship. We can, for instance, compare the following numerals:

	Ancient Tibetan:	Cantonese:
2	<i>gnyis</i>	<i>i</i> (Shanghai <i>nyi</i> )
3	<i>gsum</i>	<i>sam</i>
4	<i>bži</i>	<i>sī</i>
5	<i>lnga</i>	<i>ng</i>
6	<i>drug</i>	<i>luk</i>
8	<i>brgyad</i>	<i>pāt</i>
9	<i>dgu</i>	<i>kau</i>

Anyone can see that these words are really identical, offshoots of the same primitive forms. And we can at once make an interesting observation. Consonant groups, not allowed in modern Chinese, not only occur in Tibetan, but are a striking feature of it. In this language we find most of what we lack in Chinese, and what we expect to find in a less obtruncated language. Inflections are common. Thus, in substantives we find quite a number of inflections, and, be it observed, these endings are not, as the Chinese auxiliary hangers-on, occasional and optional, but obligatory:

Nominative	<i>lus</i> 'body'
Genitive	<i>lus-kyi</i>
Instrumental	<i>lus-kyis</i>
Dative	<i>lus-la</i>
Locative	<i>lus-na</i>
Ablative	<i>lus-nas</i>
Terminative	<i>lus-su</i>

In the verbs, again, we find a wealth of inflections. Tense stems are formed by variations in the sound of the word: to the present stem *gtong*, 'gives', belongs the perfect stem *btang*, 'has given', the future stem *gtang*, 'shall give', and the imperative stem *t'ong*, 'give'. To a great extent the endings show the function of the verbal form. Thus, the infinitive has *-pa* or *-ba*; *gtong-ba*, 'to give'; the gerund, *-te* or *-ste*; *gtong-ste*, 'giving', *btang-ste*, 'having given', etc. We see that the characteristic qualities of modern Chinese are scarcely to be found in its more archaic relative, Tibetan; neither the great simplicity of its phonology, the result of an antipathy to consonant groups, nor its «isolating» nature, i. e., the absence of inflections, of morphology.

But the two have one important feature in common, monosyllabism, simple words consisting of a single syllable. And we run no risk in assuming that this is very old. The original Sinitic language may very well have possessed dissyllabic or polysyllabic primary words, and some scholars are even prepared to admit that there are traces of them in the oldest form of Chinese; but broadly speaking we may venture to assert that Chinese monosyllabism must go back to remote antiquity. Tibetan and Chinese must have parted company long before the birth of Christ. Throughout the Chinese annals of the first thousand years B. C., the Tibetan tribes are spoken of as foreigners, as barbarians. The monosyllabism the two languages have in common, is most easily accounted for as a feature dating from the time when the two tongues were not yet divided. We are not compelled, of course, to account for it in this way, for it may be due to two parallel developments, in that both Tibetan and Chinese

may have contracted dissyllables into monosyllables in historical times independently of each other. But there are other indications that point to monosyllabism having existed in Chinese at as early a date as the beginning of the first millennium B. C. Thus, for instance, a number of songs are preserved from that period, and we might look at the first song in Shī-king, 'The Odes'.

As we do not know how the symbols in which the song has been handed down to us were read in pre-Christian times, let us read them as they would now sound in Pekinese:

*Kuan kuan tsü kiu*  
*Tsai ho chī chou*  
*Yao tiao shu nü*  
*Kün tsī hao k'iu.*

or in Cantonese:

*Kuān kuān tsü kau*  
*Tsoi ho chi chau*  
*Yu tiu shuk nü*  
*Kuān tsī hou k'au*

Without knowing the exact pronunciation of the Ancient Chinese syllables, we can nevertheless discern that the rhythm, the metre, is based on monosyllabic words. Note also that this strict rhythm occurs in dozens of poems, so it cannot be an accidental circumstance. If we were to insert a dissyllable here and there, the metre would be hopelessly broken.

On the other hand there is a good argument for assuming that the isolating nature of Chinese, the lack of terminations and derivatives, already existed to an essential degree in pre-Christian times. It is afforded by the Chinese script itself. It must have been the rigidity of the

Chinese words, their invariability and their incapacity to assume modifying additions, that made it possible for, and directly induced the ancient Chinese, when they created their script, to choose this expedient of composing a single and invariable symbol for every entire word. There were already a number of symbols of this kind in the second millennium B. C., but the vast majority of them were successively created during the first millennium, when the morphology of Chinese had already been, perhaps not entirely obliterated, but certainly much reduced. It is not surprising that the primitive script-makers, in their search for symbols for the phenomena around them, should have resorted to drawings, such as 人 (two legs) for 'man', *jên*, and 女 for 'woman', *nü*. But if the language had a number of forms based on this stem *nü*, as we have in German *Weib*, *Weibes*, *Weibe*, *Weiber*, *Weibern*, *weiblich*, *weiblicher*, *Weiblichkeit*, the clever Chinese would no doubt soon have sorted out the common element *weib-* in creating their script, and possibly have indicated it by means of the drawing 女, and then introduced other signs — no doubt some sort of phonetic symbols — for all these endings and derivative suffixes, *-e*, *-es*, *-er*, *-ern*, *-lich*, *-lichkeit*. In other words: if the old Chinese used the conventional and invariable symbol 信 (which I will not explain in this connection; it is now pronounced *sin* in Peking) for all the ideas *true*, *truth*, *trust*, *trusted*, without any modification of the symbol and without adding anything to it, it is, to say the least of it, highly probable that the word *x*, which is now pronounced *sin* and meant all that, was a monosyllabic word that did not allow of in-

flections or suffixes, and underwent no such modifications as the English stem in *true: truth*.

So we see that already in pre-Christian times the Chinese language had reached such a pitch of simplicity in sound and form that it imposed a peculiar character upon the script. This peculiar nature of the writing has in its turn had an enormous influence on the later development of the language, and is therefore deserving of a brief description.

## CHAPTER II.

From all we know at present we have reason to conclude that the Chinese script is entirely a native invention. The extensive archæological discoveries in China that have been made in the last few years by the Swedish explorer J. G. Andersson, in the course of which an extensive stone age civilization, stage after stage, in six different strata, from the 3rd and 2nd millennia B. C. has been unearthed, indicate, it is true, that a strong cultural influence from the west can be traced. The ornamentation on pottery is often so strikingly like the west Asiatic from the later stone age, such as has been excavated at Anau and Susa, that we cannot dismiss the thought of a cultural connection — not by any means necessarily the migration of tribes to China, but in any case commercial intercourse and cultural influence. But among these very comprehensive finds there is no trace of writing; it is therefore certain that the Chinese script was not borrowed from the west, even in its most primitive form, and that it is a native invention. The oldest specimens of writing we know from China are of two kinds. One is the short, primitive inscriptions on bronze objects, some of which are assigned to the 2nd millennium B. C. by Chinese experts and collec-

tors. But the grounds for assigning the dates are so defective, and expert opinion is so capricious in this case, that it tells us very little. Things are far better with the other group of material, consisting of several thousand pieces of tortoise-shell and bone, unearthed in the province of Honan in 1899. In ancient China they were used in prophesying. They were touched with some red-hot object, and the cracks that appeared were read off and interpreted by the diviners. On the pieces are inscribed the questions that were put to the spirits of ancestors to be answered, and these inscriptions are of the highest value. There is no doubt about the genuineness of a great part of these finds, and they give us interesting data, for in the inscriptions occur the names of the emperors of the Shang or Yin dynasty (1766—1122 B. C., according to the prevalent, though of course only approximately correct chronology), names already known to us through the oldest historical traditions. In these documents we thus have the oldest datable specimens of Chinese script — they must date from the end of the said period. The technique of the writing in these fragments is fairly high, so that we may assume that the earliest attempts at writing were made long before.

It goes without saying that Chinese script has gone through a long and fluctuating development between the second millennium B. C. and our own day. And this not only in technique, but in principle. For the study of this development we have copious material, and the Chinese palæographers have been an industrious race ever since the first learned palæographic work, *Shuo wên kie tsî*, was published by Hû Shên a century after the birth of Christ.

The oldest stage of Chinese script was simple drawing,

the representation of objects; and a picture of this kind represented a word as an entity. Even to-day some of these pictures are quite clear: 田 'field', 人 'man', 木 'tree', 川 'river', 工 'carpenter's square, work' etc., are still easily comprehensible in these modern forms. The following, on the other hand, ought to be seen, to be properly appreciated, in the form they had about 200 B. C. (the so-called «small seal» style): 日 'sun' (now 日), 月 'moon' (月), 子 'child' (子), 皿 'sacrificial bowl' (皿). And when we get to the small seal forms of 鹿 'deer' (now 鹿) and 角 'horn' (角), it is necessary to go still further back. The Honan fragments have 𠂔 and 𠂕. And if the small seal style has 𡚦 for 女 *nü*, 'woman', it is scarcely better in the Honan fragments: 𡚧.

Pictographs were an excellent form of script as long as they represented concrete things, and the oldest lexicon in China, the just mentioned Shuo wên, contains 364 separate pictures. But the weakness lay in the fact that it was ill-suited to representing abstract ideas. Pictures could certainly be used for this purpose in a transferred, symbolic sense, and we can find several ingenious creations of this kind. If for the word *kiao*, with the primary meaning of 'to cross', then 'entangle, intercourse', etc., the Chinese required a pictograph, they put 𡗗 (交), a man with legs crossed. This expedient was employed to some considerable extent. Shuo wên contains 125 such symbols. But this clearly became an insufficient expedient, and new methods had to be resorted to.

Before going on to describe these, we must comment on



the more technical side of the development. It is evident that a script, the fundamental idea of which is the picture, pictography, lent itself to variation in its execution. The Honan fragments already show a motley variety in the execution of one and the same word-symbol, and things grow worse in proportion as Chinese feudalism blossomed out in the 1st millennium B.C. The Chinese dominion then stretched over the whole of what is now N. China, many were the seats of culture — the capitals of the various feudal princes, often at great distances from each other. At every such petty court there were historiographers, astrologers, and administrative scribes of different kinds; and this of course led to a lack of uniformity in the use of the script. The drawing of an object may vary considerably, and the scribes of one court developed one tradition in forming a certain symbol, another school another; thus arose a number of variants of every current word.

A contributive cause was the development of the implements used. On the Yin dynasty fragments we find the inscriptions finely carved with a sharp point. This method was soon replaced by writing with a kind of wooden pen dipped in a black fluid upon wood, bamboo, or occasionally silk. This afforded a quicker and freer manner of writing, and the individual scribes had more scope for carelessness and the tendency towards worthless variants. Tradition tells about the reactions against this, and about attempts to normalize the symbols into the so-called *ta chuan* ('great seal') style. But it was not until the fall of the feudal system, and the unification of the country under a central imperial power that the script itself was definitely reformed in a uniform direction. In 221 B. C. the prince of Ts'in had overcome all his rivals, and proclaimed himself abso-

lute monarch; shortly afterwards his minister Li Sī decreed the «small seal» form of script, *siao chuan*, which was undoubtedly based on the script in vogue in the principality of Ts'in, but with the symbols greatly reduced and abridged. This afterwards became the basis of all subsequent scripts in China. The camel's-hair brush, an invention almost contemporary with Li Sī's reform, modified the technical realization of the «small seal» characters, and resulted first in the somewhat stiff «clerical style», *li shu*, and after a few centuries in *k'ai shu*, the present elegant official form, (characterized chiefly by breaking up the old curved lines and circles into straight lines, angles and squares), — parallel with which came the cursive writing, the most extravagant forms of which, *ts'ao shu*, «grass characters», constitute a freakish stenography that is the terror of all Sinologists. Here is a specimen, the word *niao* 'bird':

Small seal


official form






cursive forms

However, it often happened that Li Sī misinterpreted the old symbols, and that in producing his own simplified symbols he led the script into devious ways. The learned men of China soon recognised this, and saw that Shou wên, which unswervingly adheres to Li Sī's «small seal», could not be the last word in palæographic science. As a consequence, diligent scholars have busied themselves through the ages in copying from old bronze inscriptions from the 1st, and in isolated cases from the 2nd millennium B. C., all sorts of variants of the forms occurring before Li Sī, and have exposed many of his mistakes. The collections of material thus published are quite voluminous, and they are all the more valuable because most

of the objects from which the forms were copied have in the course of time vanished into unknown space, or been destroyed in the wars and revolutions. Of special value are of course the last acquisitions, the inscribed bones and pieces of tortoise-shell from Honan, dating from the 2nd millennium B. C. There we have an abundant choice of signs which we know for certain to be older than the Chinese collectors' many variants from the first millennium, and these new signs have made it possible for a group of eminent scholars in China, Lo Chen-yü, Wang Kuo-wei and others, and for their English colleague, L. C. Hopkins, to lead Chinese palæography a considerable step further, in many and important points, than it had been brought by Chinese palæographers in the past centuries. Just a few examples:

We have two modern signs 大 *ta* 'large' and 天 *t'ien* 'heaven', which ever since Shuo wên have been regarded as associated. 大 was originally the sign for a (grown-up) person, as follows from the sign 立 *li* 'to stand' (now 立), where we have the same figure standing on the ground. The prevalent interpretation of 天 *t'ien* 'heaven' is 一 what is above man 大, certainly not a very ingenious but in itself fairly acceptable idea. Now the fact is that *t'ien* 'heaven' was a religious conception of fundamental importance in the earliest China, *t'ien* and *Shang-ti* 'the emperor on high' were synonymous terms, and the old religious empire was based on the conception of *t'ien* as ruler and highest power in the universe and the emperor as the representative of heaven on earth, a holy man who had the investiture of heaven as governor of the visible world. *T'ien* was thus about synonymous

with God, and this god was certainly regarded in later stages as a somewhat abstract, impersonal being, but in the earliest ages undoubtedly as far more concrete. Palæography now shows us that *t'ien* was a purely anthropomorphic conception, for among the different variants of the sign to be found on the Honan fragments, several have a circle or clod at the top, and it is fairly obvious that  is a drawing of a person, a portrait of God.

A further example: The word *hing* 'to go' in the «small seal» is drawn thus:  (now: ); in all times it has been interpreted as 'footprint', an interpretation as simple as it is acceptable. In the oldest documents, however, we find a form that differs slightly from this drawing;  and the deviation, though slight, is fraught with momentous consequences, being perhaps fatal to the interpretation 'footprint'. The modern Chinese palæographers rightly point out that this cannot possibly be a drawing of a footprint, and that it must be otherwise explained. In pre-Christian texts *hing* not only means 'to go' but also 'where you go, foot-walk, road,' (especially in the palace of the king or some prince, say the learned), and  is a very striking drawing for a street-crossing. For the primitive inventor of symbols it was of course easier to draw *hing* 'road' than the abstract *hing* 'to go'; hence the sign. The discovery of the new form scarcely furnishes conclusive proof, but in any case it gives food for reflection, and shakes one's faith in the traditional Shuo wên interpretation of  as 'footprint'. But investigators must keep their heads clear in this matter. If the Honan finds are a few centuries earlier than the «small seal», it by no means follows that their

symbols are invariably more original than those that are supported by documentary evidence from some centuries later. It is of course probable that 𠂔 'footprint' is a distortion, a popular etymology for 𠂔 ('road') 'go', but it is also conceivable that the opposite is the truth, i.e. that in some clerical centre, contemporary with and parallel with the Honan finds, they already wrote 𠂔, and that this was the original symbol, which was misinterpreted in Honan and altered to 𠂔, though we happen to have no confirmation of the former in the excavations. The «small seal» characters, based on the mode of writing in the Ts'in principality, may have had a long ancestry, and many cases where it seems to us now that there are misconceptions and corruptions, may some day prove to date back to very early times. So we must not make the logical blunder of regarding the Honan bone characters a priori as the oldest and authoritative forms, merely because the Honan finds are the oldest datable ones we have got to know so far. Not until systematic archæological excavations have been carried out in the capitals of all the old feudal kingdoms shall we get a clear conception of the store of characters in the various clerical centres, and then we shall finally be in a position to attempt to decide which of them sprang from common original forms, and which arose in a quite independent and parallel way.

Thus we see that the old picture-writing offers a wide and important field for the palæographer, a field in which an enormous amount of work is still to be done. And the results will be of the greatest importance, not only from the point of view of the history of the characters, but also from an archæological point of view. Valuable and often

quite unexpected light is thrown on the oldest Chinese culture by a study of the characters.

However, we must return to the question of the fundamental development of the Chinese script. So far we have dwelt only on the first stage, the pictures themselves, used now as signs for concrete phenomena, such as 日 'sun', now in a symbolic function, such as 交 'to cross'. Whatever ingenuity the old Chinese showed in creating such signs, these soon proved insufficient, and as early as the close of the 2nd millennium B. C. we have examples of the next stage: two or more of the old pictures were combined to make signs for new ideas. By putting together the sign for 日 *jī* 'sun' and 月 *yüē* 'moon' into 明 they got a character for the word *ming* 'clear, shining'. 女 *nü* 'woman' and 子 *tsī* 'child' give the new character 'good, to love'. This ready method was utilized with great success, and Shuo wên includes no less than 1167 logical combinations of this kind. (For a number of examples see my work quoted p. 10 above). They were ingenious and excellent signs for indicating words that could not be expressed by a simple drawing. But there was one great drawback about them: they were too difficult to invent, and demanded too much of the inventor's ingenuity. In proportion as the script grew into something more than short notes about prophecies and administrative data, a method was called for which, without subtleties, could supply characters for thousands of words. Finally one arrived at a kind of script that was phonetic without being alphabetic. This method did not spring up all at once, as a creation of genius in a reformer's brain; it was arrived at gradually by two separate but concurrent ways.

At a very early stage one had already had recourse to a simple makeshift for indicating abstract ideas, an expedient which would have been fatal if adopted on a large scale, but which was good enough in isolated cases. This was to borrow the character for a concrete word in order to signify a homonymous abstract word. We have a word that in Mandarin is pronounced *k'iu* and means 'fur-coat', another, also pronounced *k'iu*, meaning 'to seek'. The two must have been homonymous already in very early times, for when the scribe wanted to write the abstract *k'iu* 'to seek', which was difficult to represent by means of a picture, he wrote 求 instead (originally a picture of a fur). This occurs already on the Honan bones. He could venture on this violent loan because the meanings were so widely apart that, in spite of the trespass, no misunderstanding could arise. When the reader had before him such a sentence as:

我 求 福

*wo k'iu fu*

I fur-coat happiness;

he could at once guess that since *k'iu* 'fur-coat' gives no sense, the scribe must have meant the other *k'iu* 'to seek', and the meaning would therefore be 'I seek happiness'. In a case like this the borrowing of a character is accompanied by no risk, but when the writers began to use this expedient on a big scale, which was soon the case, it had a detrimental effect upon the lucidity of the writing, and called for an improved method.

The suggestion for this improvement was derived from another peculiarity in the script. There was, for instance, a word which is now pronounced *t'ung*, meaning 'unite,

together, mix', and further 'mixture, alloy, bronze'. The character was 同 and was originally used for both the abstract meaning 'unite, mix', and the substantival 'mixture, bronze'. But one fine day a shrewd scribe proceeded to add an elucidative element 金 for the latter, substantival meaning; this addition means 'metal'; consequently 銅 *t'ung* 'mixture, bronze', or the 同 *t'ung* that is of 金 metal. This method became quite common. Of a 盧 *lu* meaning roughly 'vessel, bowl', several specialised variants were made by the addition of similar determinatives: 鑪 *lu* 'brazier' the *lu* 'vessel' made of 金 'metal'; 顱 *lu* 'the brain-pan', the *lu* 'vessel' that is in 頁 'the head', etc.

There were consequently now two distinct categories of characters: (1) 求 *k'iu* 'fur-coat' as a symbol for *k'iu* 'to seek', by simple phonetic loan, and (2) 銅 *t'ung* 'mixture, bronze' as a specialized form of 同 *t'ung* 'to mix'. In the former case the whole 求 had to serve solely as a phonetic token. In the latter case, half the character 金 | 同 was an element indicating the special shade of meaning, while the sound and the main notion were indicated by the other half. By a combination of these two methods one arrived easily and naturally at a new and ingenious method. To procedure No. 1 — borrowing a sign for the sake of the sound —, they joined the No. 2 idea — adding a determinative to make the meaning clear. Thus, if they were looking about for a character for the word *k'iu* 'jade-ball', homonymous with *k'iu* 'fur-coat', they borrowed



**求** for the sake of the sound, just as they had done for *k'iu* 'to seek', which gave them **求** 'jade-ball'; and then, to indicate that it was not a question of 'fur coat', they completed the sign with the help of of a determinative **玉** (picture of a piece of jade), and wrote **球** 'jade-ball'. This brought them to a type of written character consisting of two independent elements that supported each other, one of them determinative, or to use a better term «a significant», the other a so called «phonetic», indicating or intimating the sound. In *k'iu* 'jade-ball' the **玉** is therefore the significant, **求** the phonetic. With this method, as simple as it was effective, it was now an easy matter to create new signs by the thousand. One and the same phonetic could, in combination with various significant, be used in a whole series of new signs:

<i>fang</i> <b>方</b>	<i>fang</i> <b>坊</b>	<i>fang</i> <b>訪</b>	<i>fang</i> <b>房</b>
'square'	'district'	'to ask'	'room'

The significant are **土** 'earth', **言** 'speak', and **戶** 'door', respectively.

But, it may be objected, for this method to be possible, there must be whole series of homophonous words, *fang—fang—fang*, etc. But this is not quite right. After inventing the method, the old writers did not insist on an absolute identity of sound, but only great resemblance. Thus, we have the word **古** *ku* 'ancient', which is the phonetic in **枯** *k'u* 'wither, decay', where **木** 'tree' is the significant. The sounds *ku* and *k'u* are the Mandarin pronunciation. We do not know, of course, how they were pronounced in pre-Christian times, but so much is clear, viz. that

古 and 枯 cannot have been perfectly homonymous, for then they could not have differed in modern Pekinese. They must have been, not absolutely identical, but very closely resembling, since 古 was taken as the phonetic in 枯. 古 is therefore in this case not an effective sound-indicator, but only a sound-intimator. The sign 枯 must be analysed thus: a word that has to do with 木 'tree' and that reminds us in sound of 古 *ku* 'ancient'; ergo, this must be *k'u* 'to decay'.

It stands to reason that this method could never have been invented or adopted but for the peculiar nature of the Chinese language, as I have emphasized above. Monosyllabism, lack of inflection, scarceness of consonant-groups in the words, limited permissible final consonant-sounds; all this made the words so equiform, and created such large groups of short words that closely resembled each other, and rimed together by the dozen, that this method of writing was not only possible, but presented itself as the most natural and the least difficult. As early as the 2nd millennium B. C. (for examples of this category we need only turn to the Honan fragments), it had been invented, and nine-tenths of all the present characters have been created in this manner.

All this goes to prove that when the Chinaman, in quite remote times, found his picture-writing to be insufficient, and turned to phonetic writing, he adopted methods that are quite strange to us Westerners. We look upon the alphabetical, analysing script as something very simple and natural, the system to write a word by successively adding one sound-symbol after another, analytically, as in

c-a-t. This system, though it may seem simple, is one of the most ingenious ever evolved by humanity, and it is entirely foreign to the Chinese mentality. From the first beginnings of picture-writing, the Chinaman aimed at making the symbol stand for a complete word, and conservative as ever he clung fast to this leading principle, in spite of and in face of the introduction of phonetic writing.

### CHAPTER III.

In Chapter II I have explained how the Chinese script is one of ideographs, and is not phonetic. It is true that it was to a considerable extent phonetic when the great category of characters was created (the great majority of the really current ones were created during the first millennium B. C.) which consist of a «significant» and a «phonetic». But it soon ceased to be a properly phonetic script, for it was not possible, as in a language with alphabetic spelling, to let the gradually changing pronunciation be reflected in the script, and modify the spelling accordingly. The Chinese characters were fixed once for all in their composition, and could not be changed in this respect. As long as the word for 'to inspect' 監 was pronounced *kam*, and the word for 'indigo' was pronounced *glâm*, it was all very well to write 藍 for 'indigo', the 監 *kam* being the phonetic which indicated that *glâm* was pronounced something like *kam*, while 艸 'plant' is the significant, indicating that a plant is referred to. But in later times *kam* was changed to *k'âm* > *kiam* > *kiem* > *kien* > *čien*, while *glâm* became *lam* > *lan*. To learn, then, that 艸 indicates the meaning, and *čien* the sound in *lan*, is rather nonsensical. And yet nothing can be altered in the character 藍,

for it is the time-honoured old-established symbol for the conception 'indigo'. So very soon, owing to changes in the pronunciation, the characters became merely formal, conventional symbols for conceptions.

This now began to assume critical importance in the ensuing linguistic and literary development that affected China. The children of a newer generation were quite differently situated, as regards their knowledge of the old forms of the language, than the children of countries with an alphabetic spelling. Thus, in India, a man, let us say in the 2nd century A. D., who used the form *sassū* for 'sister-in-law', could know, thanks to the alphabetic spelling in early literature, that his ancestors pronounced the same word *śvaśrūṣ* about a thousand years B. C. Anything of that kind is quite impossible in China. A Pekinese who sees the character 藍 in a 2000 year-old text, sees at once that it is symbol for the word he calls *lan* 'indigo', but he has no possibility of knowing how the writer pronounced the word 2000 years ago. At the most he can conclude that, since 監 *čien* 'to inspect' plays the part of the phonetic in the sign, *lan* 'indigo' must have had some similarity in sound with what has now become *čien*. There is no other expedient for him than to read the signs in the old texts as they are pronounced to-day in his dialect. A Pekinese will read them in his manner, a Cantonese in his, a Shanghai in his, and so forth.

What does this involve? Why, that the written language assumes a very different and more independent position than in our countries, where the written language is merely a fairly faithful reproduction of the spoken words, the sounds and sentences as they are pronounced.

As the phonology of the language changes, so does the written form of it and its literature, and in this way we get an ever changing phonetic garb for the words in the course of centuries. It is true that, as in Sanskrit and Latin, we have languages that live on artificially after they have died organically, i. e. after their modern representatives have become quite different languages through linguistic development. But when to-day we do a piece of writing in Latin, this only means that we clothe our thoughts in a language form as we know it to have existed in olden times in Italy. Matters are entirely different in China. There the written literature is not in the same sense a photograph of the spoken language. The materials of which the literary sentence is built, are not sound-symbols grouped so as to reproduce the spoken words, but conventional symbols for ideas; and the question how they were pronounced in olden times, or how they are pronounced to-day by this Chinaman or that, is of secondary importance. The literature is a product for the eye, and not for the ear and tongue, as a spoken language. It lives its own life as a kind of independent phenomenon that is parallel with the spoken language. And, be it noted, this has been so throughout perhaps the whole of our Christian era. For the great mass of the practically current characters were already established in their composition, even if not in their technical form, in pre-Christian times; but at the beginning of our era the language had already gone through a considerable amount of phonetic change — a point to which I shall return in detail further on.

Now one would argue that, even if the pronunciation of the words in the literary language was, so to speak, a se-

condary matter in a Chinaman's eyes — not only contemporary texts but also older ones being pronounced at every epoch as the words sounded in one's own dialect, *faute de mieux* — yet even the literary language was a language i. e. it consisted of sentences, which must have subject and predicate and various adjuncts, like a spoken tongue; we should therefore expect every literary man to have written his texts with the grammar of his own colloquial language, even if the nature of the script prevented him from writing down these sounds so that they could be exactly reproduced from the script by men from other dialect areas, or by a later generation. But not even this assumption holds good — once again, owing to the script. Let us look into the circumstances.

We have seen how the Chinese words in the oldest stage must have had a considerably greater variation in sound than now. They had a number of consonant groups, and words could end in very many more consonant sounds than is possible now, not only in Peking, where only *-n* and *-ng* are allowed, but also in Canton, where they allow in addition *-m*, *-p*, *-t*, *-k*. It was therefore possible to produce far more dissimilar sounding monosyllables than is the case to-day, with the modern narrow rules, and therefore, too, homonyms — such as *bear* 'ursus' and *bear* 'nascor' — were not especially numerous; possibly not more so than in our western languages. It was consequently possible in remote times to build up sentences composed of these simple words, and yet be understood in pronouncing them. No confusion arose in the mind of the listener. An examination of the literature from, for instance, the time of the Chou Dynasty (1122—256 B. C.) reveals, in fact, that the vocabulary consists mainly

of simple words of this kind. That this was not only the result of a striving to write as concisely as possible, in order to save trouble when documents had to be drawn up by means of a primitive technique, but that these old texts are the natural reproduction of the spoken language, is clear from the fact that the language is just as short and concise, with simple words, even when animated conversations, or philosophical discussions were reproduced. But when the sounds were cut down and simplified, when consonant groups became simple consonants, e. g. *glâm* > *lâm*, when final consonants were dropped, so that *p'ag* became *p'a*, *ljid* > *lji*, and when later in N. Chinese, contrary to what was the case in Canton, even *-p*, *-t*, *-k* were lost, so that *lit* > *li*, *lip* > *li*, *lik* > *li*, masses of words that had once been clearly distinguished, gradually became identical in pronunciation, so that a vast number of homophones arose. Thus, in Pekinese we have a 衣 *i* meaning 'garment', a 一 *i* meaning 'one', a 揖 *i* meaning 'bow, salutation', — words that were once quite different in pronunciation, as can be seen from the Cantonese forms, where 'garment' is *i*, 'one' is *yăt*, and 'bow' is *yăp*. In a small Pekinese lexicon with about 4300 simple words there are 8 characters (word signs) that are all to be read as *i* (even tone), 17 to be read as *i'* (rising tone), 7 to be read as *i* (broken tone), and 37 to be read as *i* (falling tone). There are 6 *yen*, 13 *yen*, 8 *yen*, 15 *yen*, etc.

As I have said, the pruning and simplification of sounds that have given rise to this mass of homophones in Pekinese, and similar conditions in the other dialects, began very early, and already in the 6th century A. D. — a period whose sound-system it has been possible to reconstruct, as will be pointed out later —, there were a great



number of homophonous words. Now what did this involve? In part that if one pronounced the old literature, consisting chiefly of simple monosyllables and a few auxiliary words, with the only pronunciation that was known, i. e. one's own dialect pronunciation, it became unintelligible to the ear, which it had not been when it was written, and when the words were still sufficiently unlike each other. And also, and especially, it involved that if one were to speak with the same simple word-material and the same simple grammar, it became unintelligible to the person spoken to. A few centuries after Christ, (the exact date is impossible to fix), the spoken language had to procure new expedients which would make the language clear and practicable, in as far as the pruning of words had caused them to become identical and undistinguishable, i. e. incomprehensible to the person addressed. And this gradually took place. The vocabulary was largely changed into compound words instead of simple ones, and this was a great step towards making things clearer.

Where before they had said 見 *kien* 'to see', they now like to say 看 見 *k'an-kien* 'look-see', two synonymous words joined. Where before they had said 意 *i* 'meaning', they now say 意思 *i-sī* 'meaning-thought'. And grammatically they got a number of new resources. Where formerly they had expressed themselves so 七 人 (now read *ts'i jên*) 'seven man', in later times they said 七 箇 人 *ts'i ko jên* 'seven piece man', 'seven individuals men', and by this addition of a grammatical determinative they indicated that the following word was a noun, a

piecemeal, individually occurring phenomenon. If formerly they said 一石 (now read *i shī*) 'one stone', they now say *i-ko shī-fou* 'one piece stone head'. The word *fou* 'head, end' became a suffix indicating knobby objects. I will not here describe in detail all the various grammatical devices that have been evolved to make up for the obscurity caused by the pruning of sounds. Suffice it to say that evidently by the end of the first millennium A. D. there was a spoken language which, in the matter of vocabulary, had compound words to a very great extent where the pre-Christian language had simple ones, and with quite a number of grammatical phenomena to which the old language was a stranger. In proportion as the various dialects have branched out, they have of course also diverged in this respect, so that the grammatical auxiliaries are not the same in different parts of China. But they all resemble each other in this respect that they have developed a number of novelties with regard to vocabulary and grammar.

Here again it is the curious nature of the Chinese script that has brought about a number of linguistic phenomena peculiar to China. If there had been an alphabetic script, based on and reproducing the spoken language, there would have been two paths to choose between. One could either have continued to write the language of one's fathers, with words spelt as they had been spelt many centuries before, and with the grammar of this old tongue. This would have resulted in a classical literary language, in contradistinction to the spoken language, just as the Indians continue to write in Sanskrit even when this has broken up into various descendent languages, and as the Italians continued for a long time to write in Latin. Or, again,

one could have let the changes in vocabulary and grammar show themselves even in the written language. Just as the vocabulary and grammar of Old English had to give way little by little in English literature to those of New English, as the language changed, so might have been the case here. But in China the ideography led in some respects into other channels. Even with the script such as it was, it would have been quite imaginable to let the written language follow the development of the spoken language, — not in respect of the pronunciation of the words, but of the vocabulary and grammar. There was certainly nothing to stand in the way, after one had proceeded from saying 一石 (however that was pronounced) 'one stone' to saying *i-ko shī-t'ou* 'one piece stone head', of also following it up and making the written form in literature 一箇石頭. And this was what actually happened, though only in an inconsiderable part of the literature, i. e. a novel and romance literature which was of a popular nature and which was regarded as valueless entertainment, and a very limited dramatic literature. In the great mass of written work, however, the scribes continued to write in a far simpler and more curtailed vocabulary than the spoken language could, for the sake of lucidity, get on with, and the grammar remained the old pre-Christian grammar, with insignificant changes. This worked excellently, for the obscurities that would have arisen in speaking with such a vocabulary and grammar, never threatened, since the script, being ideographic, precluded any misconceptions. In Pekinese it is absolutely impossible to decide, when one says *i*, whether he means 'costume' or 'one' or 'salutation' or any other of

the various *i*-'s. If you wish to make it clear that you mean 'costume' you must therefore say *i-shang* 'costume-costume'. This is quite unnecessary in the script, for 衣 is an unchangeable and unmistakable symbol for *i*- 'costume', whereas *i*- 'one' is written 一, and *i*- 'salutation' is written 揖. Thus, though the script did not allow the Chinese to write their old language with the retention of the form of the words and the spelling, as one could in India, it allowed them nevertheless to continue using the same short vocables and the same grammar; in a word, to go on writing Old Chinese as a literary language, but an Old Chinese that was only for the eye, and was inarticulate to the writer himself, since the old pronunciation was unknown, and since, if read aloud with his own rendering of the sounds, it was incomprehensible to the ear without the help of the eye, owing to the great number of short and simple homophones.

The reason why the old language was retained, was of course the irresistible power of tradition in China. The Confucian literature, with occasional documents dating from the 2nd millennium B. C., and its centre of gravity falling in the middle of the 1st millennium B. C., enjoys a sanctity and veneration only comparable with that of our Bible. In the middlemost centuries of the first millennium also flourished the philosophers who are regarded as the patriarchs of a great Chinese religion, Taoism. Then, a century B. C. came a mighty historical work, the famous Si-ma Ts'ien's historical records, which have been studied with incredible diligence, and have become the pattern for all the later great dynastic histories in China. With such a fund of classical literature it is

quite natural that the language stored in it should become a model for the literature of all succeeding ages.

Add to this that the short word-material gives the old language a laconicism, a lapidarian stamp, which is not only practical, in that it is concentrated and quickly written, but also makes an impression of austere pregnancy, pith, and virility, which has endeared itself to the Chinese literary connoisseur.

We have now seen how China obtained this curious literary language, which differs so much from the modern spoken language; a classical language that is quite different from its analogues in India and Europe, Sanskrit and Latin, inasmuch as its word-material is directly comprehensible to the eye through its conventional ideographs, while the pronunciation of the words is of little consequence, since nobody bothers himself about how they sounded in olden times, and everybody gets along, *faute de mieux*, by pronouncing them as they sound in his modern dialect, a procedure that is of hardly any value, because a literary sentence read in such a way cannot be grasped by the ear. This curious literary language has lived its independent life, parallel with the spoken languages, through many centuries, and is still to be seen in the great majority of documents published, not only in books but in newspapers, periodicals, official publications, etc. The degree of archaism varies; the word-material and grammar may in some styles approach closer than in others to those of the spoken language; but in the main the written language in China has a position far freer from, and more independent of, the spoken language than in our countries.

If script and speech have thus in the main become two

parallel and independent phenomena, this has of course not prevented a reciprocal influence being at work; this has sprung up in the course of time, and has been of great importance: it has given rise to a number of curious phenomena to which we will now turn our attention.

Let us turn first to the influence of the spoken language on the written. At an early stage, in pre-Christian times, before the spoken and written languages had diverged from each other into the more independent parallel existence I have just spoken of, the vocabulary of literature was of course the same as that of conversation. But a living language never stands still. Old words fall into neglect, and new words arise in a constant stream. The linguist does not perhaps admit this in toto. He finds that, apart from purely foreign borrowings that have become incorporated, and slang words that arise one day only to disappear the next (though occasionally they remain by some chance), the new contributions are not so overwhelming, since the novelties are most often mere variations of known stems, or dialect forms of words, or dialectal derivatives of stems that already occur in the language in another guise. And he finds that words do not die easily: if you dig deep enough you will always find some out-of-the-way spot where they persist. But the average speakers and writers are not as a rule linguists, and, superficially viewed, new words perpetually appear to crop out from some unknown void and be incorporated, while others are ousted and disappear. In our languages we are used to the literature of an epoch being a reflection of the vocabulary that was in use during that epoch. Of the vocables that occurred and were com-

mon in the English speech of the 17th century, and therefore in its literature, many are now antiquated and cannot be used in present-day English without producing a ridiculous effect. In Chinese the case is quite different. A word once introduced there and used in good literature, is thereby, owing to the Chinaman's devotion to that literature, immortal. The best of the old texts are learnt by heart by all the students of one century after another, and their vocabulary is stored in the consciousness of the educated like golden grain. A new author unhesitatingly makes use of vocables that were in use long before the Christian period, mingled with vocables that only came into use many centuries later. Chinese literature is thus an enormous store-house, where everything that has once been admitted is faithfully looked after and used; nothing is too worn out to be utilized, and perpetually new additions create a veritable *embarras de richesse*.

If we follow the development of Chinese literature we shall find that the stock of primary words (uncompounded words) was increased at an enormous rate in the centuries immediately before and after the birth of Christ. From being only a few thousand (in the preserved literature) before 200 B. C., the number of simple words — provided with signs on the principle of one significant and one phonetic — was many times doubled during this period. In later times, when the old literature had already got such a nimbus that one did not venture to go beyond the compass of its vocabulary, this new-creating of symbols ceased altogether, the literary new-formations being confined to new combinations of the old primary words; and a number of popular words that arose in the modern spoken tongues therefore entirely lack symbols; if a new

symbol was created for a new word which it was practically necessary to write, this symbol was stamped as «vulgar», and wo to the scribe who dared to introduce it into a serious script.

The problem of explaining the origin of the mass of primary words that are disclosed by the innumerable new characters created in the centuries immediately before and after the birth of Christ, is one that will give philologists great trouble. The largest Chinese lexicon includes close upon 50,000 different characters, and even if they are to a great extent mere variants, the fact remains that a Chinese-European dictionary which would include the most important words in Chinese literature, would have to put in at least 12—14,000 symbols for different primary words. It is true that we are unable to give the exact pronunciation of these many words in their old style, but we have already mentioned that there is good reason to assume that, as far back as the time of their appearance in literature, Chinese was monosyllabic, and lacked to a great extent the faculty of varying the stems of words by adding prefixes or suffixes. How then are we to understand this crowd of new primary words? Many of course existed long before, and only received symbols later, while a smaller number are probably loan-words from neighbouring peoples or conquered tribes. But there remain huge series of pure synonyms, apparently primary words independent of each other.

The occurrence of these numerous primary words with their own symbols, is no doubt due in part to the rapidly progressing shifting of meanings and specialization of meanings that takes place in every language. Assume that in Ancient Chinese there were four words all pronounced, let



us say, *γ<sup>w</sup>an*, meaning respectively 'return', 'bracelet', 'top knot (of hair)', and 'palace wall'. These must have been regarded by the old scribes as four independent words, and therefore each of them got its symbol: 還, 環, 鬟, 寰.

But as we can see, the scribes have included the same element in each of them (𠄎), which seems to indicate that in this particular case they actually were sensible of the etymological connection between the words; they are only specialized by different significants: 辵 'go', 玉 'jade', 影 'hair', 山 'house'. The stem *γ<sup>w</sup>an* evidently had a fundamental signification of 'ring, round, go round, turn', from which, in the course of time, the widely different conceptions 'return, bracelet, top knot, palace wall' were evolved. In other cases they did not even surmise the connection of the words, and wrote them with entirely different symbols. Thus we have 集 Pekinese *tsi* (Anc. Chinese *dz'iap*) 'to gather, collect, union', and 輯 Pekinese *tsi* (Anc. Chinese *dz'iap*) 'harmony', etymologically identical but regarded by the Chinese as two independent words. It will therefore be an extremely difficult task, as I have said, for etymologists to reduce, so to speak, the great number of primary words in Chinese to a reasonable and genuine number, and to identify those that are merely sense variants of etymologically one and the same word. The phenomenon here touched on also explains — which is only another aspect of the same problem — how it is that we can have a multitude of words differing in sound as well as in symbol, for one and the same idea; or, to put it in other words, how we can have so many synonyms, when other languages have

considerably fewer. Assume, for instance, that a word for 'wall' (palace wall) 寰  $\gamma^{w'an}$  (I give here the ancient pronunciation of the words discussed), is etymologically the same word as 還  $\gamma^{w'an}$  'turn round'; that another 'wall' (protecting-wall, screen) 屏  $b'ieng$  is really the same word as a 駢  $b'ieng$  which means 'to put in a row' — thus alluding to the stakes in a palisade; that a third 'wall' (low wall) 垣  $j_i^{w'vn}$  is etymologically identical with 園  $j_i^{w'vn}$  'enclosure', 'garden'; that a fourth 'wall' (outer city wall) 郭  $kuák$  is the same word as 櫛  $kuák$  'outer coffin' (the fundamental sense of both being 'outer cover'); that a fifth 'wall' (rampart) 堡  $páu$  is etymologically identical with 保  $páu$  'protect, guard', and 褓  $páu$  'swaddling clothes' («protection»); — it is now easy to understand how one obtained such a multitude of apparently primary words for one and the same idea. They demand, what was actually the fact in China, a sufficiently long linguistic development in an advanced stage of culture for a far-reaching alteration in meaning, and a specializing of the word-material to have been able to take place.

If consequently the spoken language for many centuries powerfully affected the written language by continually supplying it with fresh stores of apparently primary words, while the old supply of primary words remained living, was utilized, and honoured, resulting in a vocabulary that was immense and abounding in synonyms, the literary language has, in its turn, affected the spoken language, or rather spoken languages, in a curious way. From our western languages we are accustomed to a strong influence emanating from a common standard language, a language that

is the regular written medium, and acting on the dialects. Through schooling, by the reading of books, and especially newspapers, the masses get their vocabulary increased by a quantity of words that do not really belong to their respective dialects, words of a learned character, abstract words, etc., as well as synonyms to everyday words that do exist in their dialects. A man will speak of *lasses* and *lads*, but understands the corresponding standard expressions *girls* and *boys*; or of *fells* and *burns*, for the standard *hills* and *brooks*. He disdains the latter words, but he knows them, and will perhaps actually use them in certain situations, if he is afraid of not making himself understood with the former words, which are more natural to him. He gets this new stock of words, however, from another dialect, and in the phonological garb of another dialect, for our literary languages, our standard languages, are merely a written reproduction of the dialect of some dominating area, frequently that of the capital or its environs, with a number of alloys and additions, it is true, yet in the main with the character of such a living linguistic variety. In China the case is entirely different. The spoken language of the capital has, to be sure, a certain shimmer of being fashionable, and for practical reasons it is to some extent adopted by the world of officials, but a literary role, such as it plays in western countries, it has never played as yet — we will return later to tendencies in the opposite direction in quite recent times. It has scarcely happened that a word from the idiom of the capital has passed, in its Pekinese phonetic form, into other dialects through literary influences. China has, as I have said, a number of dialects, but the word dialect must not be taken in the commonest sense

of a provincial mode of speech as against a common educated form of speech and writing, a standard form of the language. The Chinese dialects are independent, parallel idioms, which make up the Chinese linguistic unit, just as did the dialects of Old Greece, Attic, Doric, Ionian, etc. None of them has been raised to the position of a literary language that dominates and influences the others. The literary language is the old language, the written language, appealing to the eye rather than to the ear, and its effect on the various idioms, or dialects, (as I take the liberty of calling them, with the above reservation), is not produced through the ear, their words being borrowed by the dialects in the phonetic form they have in any dominating dialect, but direct from the books to the speakers of the various dialects. This is due to the character of education in old China right down to the beginning of our century. Primary education was given in the simple little village schools, and for higher studies people had to do the best they could for themselves, either with the help of books alone, or with the help of some capable local private teacher. The part played by the authorities was in China merely the final control of the results of studies, i. e. examinations, always in writing. There has therefore been no influence through the public educational system which could engraft the pronunciation of one locality on the students of another; in each locality instruction went on isolated, with apostolic succession from teacher to pupil, entirely confined in pronunciation to that of the particular dialect, even in dealing with the most classical old texts. Words in these texts which were such as are still in common daily use in the dialect in question, were naturally pronounced as they were spoken in that dialect — the ideographical script

does not invite to any other mode of pronunciation — and no student has bothered himself in the least about how the words were pronounced in other parts of the Empire. If in the texts they came across words that were no longer in use in that dialect — which would of course often happen — they never asked themselves in their studies how these words were pronounced in Peking or Canton, etc., because there is no authoritative standard pronunciation; instead, they consulted their written helps. In the old lexicons you look up a word according to purely graphic points of appui — the words are arranged there according to their significant (meaning-elements) and the number of strokes in the character. There you will find a direction that the word *x*, the meaning and pronunciation of which were unknown to you, means so and so and is pronounced like the word *a*, a word fully current in your own dialect; and then you learn the new word with this pronunciation. A concrete example: In the classical language the word for the notion 'to say' is this: 曰. It is now out of use, and replaced in the spoken languages of N. China by the word 說 (Pek. *shuo*). The student looks up the symbol 曰 and finds that it is pronounced like the word 越, which means, inter alia, 'the more'. This is fully current in the modern dialects. In Peking 'the more' is *yüe*, and a Pekinese therefore also pronounces 曰 *yüe*. In Canton 'the more' 越 is *üt*, and therefore a Cantonese also pronounces 曰 *üt*. In Shanghai 'the more' is *iö*, and there 曰 is accordingly pronounced *iö*. In other words: round a little kernel of two or three thousand primary words which have remained current from old

times till to-day in the dialects, the student of every district builds up a whole structure of pronunciations of literary words, with the help of the written guides. *x* is pronounced as *a*, *y* as *b*, *z* as *c*, etc. For the linguist, this involves a curious doubling of the material. On the one hand he has a really living word to work with; a Pekinese 商 *shang* 'discuss' corresponds to a Shansi *sā* and a Cantonese *shōng*, all perfectly current. On the other hand he has a number of secondary, really dead words, in which he can find exactly the same sound-equivalents: a Pekinese 觴 *shang* 'goblet' corresponds to Shansi *sā* and Cantonese *shōng* 'goblet', a word only occurring in literature. That the sound-laws, the sound-equations are the same, is due to the fact that students in different tracts have learnt the dead words from the same old lexicons, which state that 觴 'goblet' is to be read as 商 'discuss'. But I must add one or two reservations to what I have said.

If the pronunciation of literary words, which are no longer alive in a dialect, is nowadays chiefly determined by students and their teachers through literary reference books, lexicons, this has not been the case since time immemorial. Pronouncing dictionaries did not come into serious use until the 6th century A. D., prior to which the correct reading of the texts was undoubtedly handed down by learned schools through the ages, from master to pupil, the pupil learning the correct pronunciation from the master (and to a great extent memorising the texts) for, as we have pointed out, he could not make out the pronunciation from the manuscript. But, owing to the very nature of the Chinese script, which had no phonetically pre-

serving influence, this did not suffice to preserve the former pronunciation. The gradual, unnoticeable changes in the pronunciation of current words also dragged with them the words that were no longer current. When the common words of a type *-ian*, e. g. *kian*, *lian*, *mian*, etc., gradually got the *a* modified to *ä* under the influence of the *i* (*i*-mutation), and became *kiän*, *liän*, *miän*, — a change that took place so gradually that the speakers were unconscious of it, it taking several generations for *kian* to become *kiä/än* and then *kiän*, — masters and pupils just as unconsciously let even such *kian*, *lian*, *mian* forms as were purely literary words, no longer current in the spoken language, follow the stream, so that in reading the literary texts aloud they were changed into *kiän*, etc., just like the current words, there being no alphabetic writing to check the flow and recall to mind that these words were once pronounced *kian*, etc. Later, in the 6th century, when pronouncing dictionaries appeared, the pronunciation of the current words was established in them with the actual pronunciation of that time, but the pronunciation of words no longer current was established, not with their old pronunciation, but so to speak with the 6th century projection of the earlier pronunciation.

A second reservation. While in most dialects, e. g. the Mandarin dialects and Cantonese, there is no essential difference between the pronunciation of the spoken language and that of the literary language (apart from isolated words), there are certain groups of dialects in eastern and southern China, where large categories of words have two pronunciations, colloquial and literary. This is the case, for instance, in the Shanghai dialect. And the Swatow

dialect (S. coast) is so complicated that for large groups of words it even has three pronunciations (though the majority of words have only one pronouncing form). The curious usages in this dialect are nicely described by Gibson<sup>1</sup>: «There are two recognised ways of reading the written character, besides the sound of the spoken vernacular. One is usually employed in reading the text of the classics and. . . is used for no other purpose. This is called the *chia*<sup>n</sup>-*im*, or correct sound. . . . It is the sound in which a scholar would recite from memory any of the classical texts. It is also read in schools during part of the day, but when the explanation of the text is to be given, the text itself is read in another style. This other is called the *peh-im*, or common sound. It is employed in reading classical texts before expounding them, and in reading all other classes of books, letters, essays, public notifications, and all kinds of written documents. . . . Examples may be found of words or phrases which occur in three distinct forms, one in colloquial, one in ordinary reading, and one in the classical style. Thus the phrase 大學 ('the great learning') would be pronounced (1) in colloquial, *Toa-oh*, (2) in ordinary reading, *Tai-hak*, (3) in classical sound, *Ta-hiok*.»

In these dialects, with such a special literary way of reading the texts, we must not imagine that this literary style represents a pronunciation preserved from ancient times by the conservatism of the learned, as de Groot erroneously believed. As a rule, this reading style only represents an approach to the pronunciation of the Mandarin (so that the colloquial pronunciation is often more archaic), and this constitutes therefore an important ex-

<sup>1</sup>) John Gibson, A Swatow Index to the Syllabic Dictionary of Chinese by S. Wells Williams . . . 1886 p. 6.



ception to the general rule that in China one dialect has had no literary influence upon another in the matter of pronunciation. This exception, however, cannot invalidate the general rule about the phonetical independence of the dialects even in literature, and their averseness from taking impressions from each other's sound-systems.

Now in this old land of learning, China, the literary language has exercised an immense influence on the educated spoken language. Thousands upon thousands of once extinct words and phrases have been reintroduced direct from the written language into the spoken language of the educated classes as literary and elegant loans, and we can thus, thanks to this operation, witness a curious state of things. While in Swedish dialects the primary words that exist in all or most dialects (in forms resulting from the phonetic laws of these various dialects) are not very numerous — perhaps only a few thousand —, in China we have whole series of many thousand words to be found in all the dialects, identical in meaning but different in pronunciation, and with extraordinarily consistent phonetic correspondences, very strict sound-laws. Of these masses of words that are everywhere current, a very great percentage are such as have been reintroduced in the educated spoken language through influence from the old literature, chiefly in the form of compound words, standing expressions, and phrases of a literary type; and the strict consistency in the sound-equations is due to the peculiar way in which they have been reintroduced, and to which I have just referred. In fact, the whole mighty vocabulary of the literary language exists in this way potentially in every dialect, with the pronunciation demanded by the phonetic laws of that particular dialect.

China thus possesses, not one national standard language, but every important group of dialects constitutes its own centre, and each of them has all the apparatus of the cultural and technical vocabulary pronounced in its own manner, uninfluenced by other centres; a series of parallel, refined colloquial languages, like Attic, Doric, Ionian, etc., in Old Greece, each of them increased from the skeleton of a mere provincial dialect into the wealth of an educated language, rich in cultural words and abstract terms that have been borrowed direct from the old literature; each dialect has kept its own phonetic system, and has adopted, moreover, all new cultural compounds (made to the pattern of western words) in a phonetic garb which results from the pronunciation of the components in the dialect in question. There is, for instance, an expression 博聞 *po-wên* 'wide-hear' i. e. 'of wide experience (especially in literature), well-read, well-educated'. Both 博 'wide' and 聞 'hear' are purely literary old words that do not exist in every-day colloquial language (in Mandarin 'wide' is *kuang* and 'hear' *f'ing*). Through literary influence the expression exists in the educated spoken language throughout China, and is pronounced *po-wên* in Peking, *pǎ-wêng* in T'ai-yüan-fu, *pok-mǎn* in Canton, etc. In the same way, when in our day a word is coined for 'anatomy' 解剖學 ('unloose-split-science'), it is pronounced *kie-p'ou-hūe* in Peking, and *kai-p'au-hok* in Canton.

We have now touched on a whole series of deviations from what we are accustomed to in our own languages, all ultimately to be traced to the early pruning into monosyllabism and lack of morphology. For it was this linguistic uniformity that led to the peculiarity of the script, the

peculiar script led to the literary language, which is independent of the spoken language and appeals essentially to the eye. This artificial, independent nature led to continual inpourings of apparently fresh word-material, without the older material becoming obsolete, as in a natural language. And finally, this syncretism in the literary language has in its turn influenced the modern Chinese dialects.

## CHAPTER IV.

The great peculiarity in the Chinese world of letters, which has been described in detail in the previous chapter, viz., that there is an extensive classical literature which can be read and understood without the reader having a notion how the words were pronounced when the works were composed, may appear to have both its advantages and its disadvantages. To the layman it may seem convenient not to have to bother oneself about the varying pronunciations of a word during different epochs that extend over millennia. If they had alphabetic writing in China, reproducing the pronunciation of the different stages, one would be forced to learn long lists of the forms in which one and the same word was pronounced in the lapse of centuries, whereas now one can learn once for all the character, and at most its Pekinese pronunciation, and then not trouble about the phonetic fate of the word through the ages. The practical philologist may here appear to be far more independent of linguistics than in other languages, and this has without doubt been the prime reason why the Sinologist has put off as long as possible subjecting Chinese to a detailed linguistic analysis, always apart, of course, from pure lexicography. The Chinese themselves have also chiefly been exegetists, literary philologists, archæo-

logists. It is true that the learned world has been annoyed ever since the 12th century A. D. by the fact that the rimes in the oldest poetry no longer rimed, and the scholars soon realized that the language had gradually changed its pronunciation in the course of time. But the linguistic studies this gave rise to, especially in the 16th and 17th centuries, and various attempts to reconstruct the sound-system of the early language, were clumsy. They were unable to pose the problem properly, and moreover they lacked access to a quantity of material at our disposal, to which point I shall return anon. The attempts in question were therefore foredoomed to failure.

That investigators from western countries, with a few exceptions, did not devote themselves until about twenty years ago to the study of this problem, is more unpardonable. Even if we shut our eyes to the fact that the linguistic investigation of a language is a large and important field of study in itself, a task of the very greatest intrinsic importance, linguistics is an auxiliary science for Sino-logy in its entirety that possesses, not less but rather more importance, than in other philological domains. It is indispensable to the historian. The kinships and pre-history of the Chinese people cannot, of course, be determined with any certainty without the help of linguistics. But even when it is a question of actually historical periods, the historian would be hopelessly handicapped without the same source of aid. It becomes clearer every day that the history, not only of China but of all Eastern and Central Asia, lies hidden to a great extent in the treasury of Chinese literature. In that literature there exist plentiful and conscientious descriptions of the foreign races the Chinamen had come in contact with.

There are innumerable transcriptions with Chinese characters for names of nations, places, persons, gods, etc., which have to be identified, and it is impossible to do this unless we can base ourselves on the ancient pronunciation of the Chinese words. If in Jordanes the N. European names that have permitted so many interesting conclusions, had been written, not in a script that enables us to read off their pronunciation, but in hieroglyphics of unknown sound-values, his services to our early history would have been slight indeed. In China the case is quite analogous. A good example of how much more effective a key the ancient pronunciation is than the modern, is afforded by the name for Buddha, *Śākyamuni*. In Chinese this is rendered by 釋迦文 now pronounced in Peking as *šī-čia-wən*, which is a poor sound correspondence. We know, however, by ways that I shall shortly return to, that in the 6th century A. D. this was pronounced *śīāk-kīa-mīuən*, which is considerably better.

And even if we do not go beyond purely Chinese history, we must admit that it is a ridiculous state of things that, for lack of a better expedient, we read the names of Chinamen of the last four thousand years in modern Pekingese; we can confidently assert that — at least in the case of the men of the older time — their names were certainly pronounced in quite another fashion. For the historians of literature, again, the case is analogous. Not only is the old poetry quite spoilt, when with the modern pronunciation the rimes no longer rime; every form of literature loses its rhythm, its beauty, all its style, the very moment the documents cannot be read in the form in which they were created, but only in a distorted and altered form, as far as the sounds go; and thus all æsthe-

tic appreciation of a critical kind is rendered impossible. And even the most inveterate of practical philologists, who only pay attention to the subject matter of the texts, will in the long run be dependent on the assistance of the linguist. For, as we shall see, the literature of remote China is not preserved in MSS. actually dating from the time when they were first written, but they have gone through the most varied fates, and one of the most difficult problems for the Sinologist is the question of the authenticity of the oldest documents. To call upon the aid of linguistics in solving this problem, has not yet been seriously attempted, but in the long run it will be the only safe expedient. We will return to the question later.

The reconstruction of the Chinese language, as it sounded in different periods of the past, is an inevitable and central task in Sinology, a task which, among others, I have myself been occupied with for many years. A number of encouraging results have been arrived at, with methods that I will now briefly define; a huge mass of work remains to be done.

The Chinese themselves have, as already mentioned, at all times been general philologists rather than linguists. Whereas, during the first few centuries A. D. they had already become trained as textual critics, commentators, and practical philologists of great acumen, it was not until well into our Christian era that they took their first steps as real linguists, and even then with only practical objects in view.

In proportion as philological science developed, the need for practical dictionaries made itself felt. Here, two different means were adopted. In some of the dic-

tionaries the characters were arranged graphically, in others phonetically; and it is the latter that are of the greatest interest to linguists. Seeing that the script was not alphabetic, it was impossible to arrange the words alphabetically, as in our dictionaries. A different method was adopted. First, the compilers made clear to themselves the structure of the Chinese word, and it was analysed into two main parts: the first half, the initial, was only the initial consonant sound; the second half, the final, was all the rest of the word, from the first vowel sound onwards. Thus, they divided *kuo* into *k|uo*, *kang* into *k|ang*, *kien* into *k|ien*, *tsuân* into *ts|uân*, *au* into *|au*, (initial nought) etc. It was now possible, by means of two words, to give the reading of a third. If the word 干 was read *kân*, 古 *kuo* and 寒 *yân*, one could spell 干 by combining the others, taking the initial in 古 and the final in 寒: *k(uo) + (y)ân = kân*. This method of spelling was called *fan-ts'ie*, which means «turn and cut». In the old lexicons we find under every title-word its pronunciation given by *fan-ts'ie*. If we look up the word 干, we at once find: 古 寒 切 «*kuo-yân* cutting». Now it is evident that these old keys to the pronunciation are for us in modern times so many equations with algebraical values. For we do not know, from our present starting-point, how 古 or 寒 were pronounced (the sounds *kuo* and *yân*, given by me, are deduced by other means), and every such *ts'ie*, spelling, is a formula:  $x = \text{the former part of } y + \text{the latter part of } z$ . They are, however, invaluable; for by means of them we get a definite starting-point. It can be proved by cross-references that in the speech



of an ancient lexicographer, certain 20 words had the same initial *y*, certain 35 words the same final *z*, etc.

For the old-time Chinaman, who moreover knew the exact values of these symbols *x*, *y*, etc., it was easy, from this sound-analysis of the words, to arrange them in a systematic way in the lexicon. First he divided the words into three groups according to the musical accent, e. g. *kân*<sup>-</sup>, *kân*<sup>'</sup>, *kân*<sup>`</sup>, and as a fourth group, forming the «abrupt» equivalent to the former, he put *kât*. And within each such main group, e. g. all words with level tone, he arranged the words according to rime. Thus, he brought together all words ending in *-ung* into one group, then all words in *-âng* (*â* is open *o*) into one group, then all words in *-i* into one group, etc. And within each rime of this kind he put together all the words beginning with one and the same initial. Thus, within the rime *-ân* he put first 干 *kân*<sup>-</sup>, then 竿 *kân*<sup>-</sup>, 乾 *kân*<sup>-</sup>, etc.; then 單 *tân*<sup>-</sup>, followed by 丹 *tân*<sup>-</sup> and the other homophones. It is clear that a schematic arrangement like this was only made possible by the peculiar simplicity of the sounds, and the shortness and equiformity of the words dealt with.

Among the old lexicons with these fan-ts'ie and this arrangement of the sounds, the most valuable to the linguist is *Ts'ie-yün*, a lexicon which was published in 601 A. D. after long preparation, and which represents Northern Chinese in the 6th century. The work itself is lost, and it was not until the 20th century that fragments of it were found in Central Asia; but its rimes and ts'ie-spellings are preserved in later adaptations of the lexicon. These constitute one of our most valuable linguistic sources.

The arrangement in Ts'ie-yün gives us important information. I said just now that, by means of cross-references of fan-ts'ie, we can determine a series of strictly discriminated final sounds, e. g. *-âng*, *-ien*, *-uân*, etc. Now we find that Ts'ie-yün often brings together under the same rime several final sounds strictly discriminated in spelling. The rime *-âng* includes both final *-âng* and final *-<sup>w</sup>âng*, the rime *-ung* both *-ung* and *-iung* — and they undoubtedly do rime. Here again we must remember that the exact values (*-âng*, *-<sup>w</sup>âng*, *-ung*, *-iung*) are not obtained from the riming-dictionary, and must be supplied by other means. But the arrangement gives us valuable points of support, for it gives us an idea of the structure of words in the 6th century. The word is made up of initial and final. But sometimes the final may also besides the actually riming part have other elements, namely, subordinate vowels:

*k<sup>(w)</sup>âng*; *k(i)ung*.

Through these old sources, fan-ts'ie and the riming dictionaries, we have now been able to draw up the phonological groups of Ancient Chinese, i. e. of Chinese of the 6th century: these 20 words have had the same initial *x*, these 15 the final *y*, these 12 the final *z*. And the finals *y* and *z* have had the same chief vowel and final consonant, for they are arranged in the lexicon under the same rime.

When, after this, we want to get to closer quarters with the problems, we shall find valuable sources of a somewhat later date.

The condition of extreme poverty in sounds shown by all the modern dialects, is, as I have said, the result of an immemorial process of pruning and simplifying. By

the 6th century this had, it is true, already gone pretty far, but the phonological groups in the Ts'ie-yün lexicon show that the language of that time had a considerably wealthier show of initials and finals than any present dialect possesses. But the simplifying process was in full swing, and philologists of the 11th century could only with difficulty find their way about in the lexicons based on the richer sound-system of the 6th century. They had difficulty in finding the words. They then set to work to produce a systematic key, based on the language of the 11th century, to the lexicon of the 6th century. It was by imperial edict that the distinguished philologist and politician Sī-ma Kuang, a scion of an old imperial family, set up his tables of sounds, commonly called rime-tables. These became popular. They are to be found in a later adaptation in the introduction to the imperial lexicon K'ang-hi tsī-tien, published in 1716, and are thus in the hands of all students. In a series of tables are given a list of all the dissimilarly sounded syllables that Sī-ma's language possessed. Here is a specimen from one of the tables:

Ancient  
rimes

*m*   *b<sup>c</sup>*   *p<sup>c</sup>*   *p*   *n*   *d<sup>c</sup>*   *t<sup>c</sup>*   *t*   *ng*   *g<sup>c</sup>*   *k<sup>c</sup>*   *k*

寒							單			看	干	<i>ân<sup>-</sup> 9</i>
早											筭	<i>ân' 10</i>
翰											肝	<i>ân' 11</i>
曷					達						葛	<i>ât 12</i>
												<i>an<sup>-</sup> 13</i>
								眼				<i>an' 14</i>
												<i>an' 15</i>
												<i>at 16</i>
												<i>ian<sup>-</sup></i>
												<i>ian'</i>
				變								<i>ian'</i>
												<i>iât</i>

In every vertical row all the words have the same initial; in every horizontal row they have all the same 11th century final. I have not filled in all the characters in the squares, but only a few examples. Wherever a syllable, combined of a certain initial and a certain final, was in Sī-ma's language represented by an actual word, this is inserted (or, if there were several homophonous words, one of them was inserted) in the square where these columns meet. Ancient 眼 *ngan'* 'eye', 達 *d'ât* 'reach', 變 *piân'* 'change' will therefore occur in the tables in the way indicated in our specimen. To the ex-

treme left are the rime or rimes in Ts'ie-yün under which words with the final in question should be looked up. The sound-simplification between the 6th and 11th centuries having caused various old rimes to become identical in pronunciation, there are often two or more rimes in the rime column to the left. The vertical columns were arranged, as we see, according to the phonetic affinity of the initials: the gutturals *k-*, *k'*-, *g*'-, *ng*- form one group, the dentals *t-*, *t'*-, *d*'-, *n*- form another, etc. In the same way the horizontal columns are arranged according to the phonetic affinity of the finals. In every division there is the same rime except for the tones (*-dt* was regarded as the abrupt correspondence to the more lengthened *-ân*, *-ân'*, *-ân'*), and in the same great scheme are brought together rimes that are closely allied. Every square, as I have said, contains an exemplifying Chinese character, provided a word existed that was made up of the initial and final of the square in question.

Here I must again emphasize that the phonetic values I have placed in the upper margin and right margin are reconstructed in a way to be mentioned presently. Of course they do not occur in Si-ma's tables, which only contain squares filled with characters, and columns with the old rimes to the left. Though the tables make up a systematic exposé of Si-ma's language, they are in themselves a closed world. To be able to fill in the sound-values represented by the different columns, we must turn to other sources. Thanks to the rime-tables, however, we know considerably more than we could learn from fan-ts'ie and the rimes that are but inorganically arranged in the lexicon. We know that the initials 1, 2, 3, 4 form a phonetically coherent group, e. g.

gutturals; 5, 6, 7, 8 another, e. g. dentals. And what is more, we know that

1 stands to 5  
as 2 » to 6  
as 3 » to 7, etc.

If we can prove that 1 was *k*-, 2 was *k'*-, and 3 *g'*-, then we are practically certain that 5 was *t*-, 6 *t'*- and 7 *d'*-. As regards the finals we know that 9, 10, 11, 12 were identical but for the tones; if 9 was an *-ân*, the others were *-ân'*, *-ân'*, *-ât*. Likewise with 13, 14, 15, 16. And we also know there must have been close phonetic connection between 9, 10, 11, 12, on the one hand, and 13, 14, 15, 16, on the other, since they occur in the same table. In other words: we have now reached the point where we have got the words in the old language classed into sound-groups so very exactly that a result with regard to the concrete sound values which has been obtained from other sources with absolute reliability on one point, can immediately be extended to apply to a whole big category of words.

Now, what are the sources that are finally to supply us with the concrete values to be inserted in the old tables with all their *x*- and *y*- columns? We might expect valuable information from European visitors in China in olden times. Thus, for instance, Marco Polo, the Venetian merchant, undertook his famous expedition to the East in the 13th century, was for a long time an honoured and trusted man at the court of Kublai Khan in the brilliant Khan-baluk, i. e. what is now Peking, and his memoirs contain much information about, and names of, things he had seen. These names are, however, so unmethodically and primitively reproduced that they give us little help. The

same holds good of the writings of the early Jesuit missionaries in the 17th century; and here, moreover, we have reached such an advanced epoch that the material is of little value.

The real main sources of help are two: first, the modern dialects, the fountain-head; and secondly, certain reflections of Ancient Chinese in or through languages abroad.

As we know, China has two mighty river-systems, that of the Hoang-ho in the north, and that of the Yang-ts'i-kiang in the centre of the country. Both rivers are surrounded in the west by mountains, in their middle and lower course by large plains. The level country is separated from the south coast of China by long inaccessible mountain ridges, comparatively close to the sea. The whole of the huge country from the north right to the Yang-ts'i-kiang, and largely south of that, has a tolerably uniform colloquial language, which with a translation of the Chinese term «kuan-hua» or «language of officials» is generally called the Mandarin language, or simply Mandarin. As a matter of fact it is made up of a great number of subdialects, which do not, however, differ from one another so much that their representatives cannot, at least after a little practice, understand each other tolerably well. The most marked peculiarities are to be found in a little mountainous province, Shansi, in the north of China. Records from this district have hitherto on the whole been lacking, but some comprehensive records from a dozen Shansi dialects that I was in a position to make (see my *Etudes sur la Phonologie Chinoise*) have proved to be of greater linguistic value than any other Mandarin dialect. The most fashionable variant of Mandarin is that of Peking, upon which is based most of the European lexicons.

The coast regions in the south and south-east of China which are isolated by the above-mentioned mountain chains, have diverged greatly from the rest of China in their linguistic development. In many respects they are more old-fashioned than Mandarin, and they are of the greater value to the linguist in that they are almost as different from one another as they are from that. They fall into three widely differentiated groups, from each of which there are good lexicons dealing with two or three varieties that stand to each other in the relation of independent dialects.

As a second source I mentioned the reflections of Ancient Chinese in or through languages abroad.

Here we must first call attention to the great number of transcriptions of Indian and Central Asiatic words to be found in the old Chinese literature. They are in Chinese characters, and thus one gets a fair idea of how these characters must have been pronounced in order to serve as transcriptions of the foreign words. And vice versa, there are quite a number of examples of Chinese words having been written in Central Asiatic documents in the alphabetic script of the strangers, which is of course exceedingly valuable material. It would not, it is true, be advisable to build up a complete interpretative system of Ancient Chinese on the strength of such transcriptions, but they constitute an excellent touchstone by which we can control whether we have reached correct results by other means.

Far more important, however, is the reflection of Ancient Chinese in actual loan-words on a large scale in certain foreign languages. At all times China has been the great cultural source for East Asia, and when Korea, Japan, and



Annam borrowed Chinese culture *en bloc*, thousands of Chinese words found their way into the languages of these countries. No proofs have so far been brought forward to show any relationship between Japanese or Korean, on the one hand, and Chinese, on the other; and whether Annamese is allied to Chinese is uncertain. But in historical times, as we have said, a whole army of Chinese words has entered each of these languages, comparable with the importation of Low German words into the Scandinavian languages during the Middle Ages. These elements of vocabulary in the said languages are commonly called Sino-Korean, Sino-Japanese, and Sino-Annamese, respectively.

What is of such choice interest for linguists in these three varieties of Chinese, is that the loans were made long ago, that the connections were severed, and that the loan-material was preserved in the foreign countries uninfluenced by later sound-developments in China. Hence, this word-material often gives us the most valuable information, in spite of the ingrafted words having in this or that respect developed according to the sound-laws of the foreign tongues.

Least valuable of the three is Sino-Annamese. Not only is it comparatively young, dating in the main from the 9th century, but the Annamese script is of such a nature that it is hard to decide which phenomena in Sino-Annamese go back to the period when they were borrowed, and which have developed on Annamese ground.

Better in these respects are Sino-Korean and Sino-Japanese. Both are several centuries older than Sino-Annamese (Sino-Korean c. 600 A. D., Sino-Japanese 5—8th centuries), and both Korea and Japan have for a good

long time had phonetic spelling, and fortunately for the linguist no spelling reformer has dared to extend to them his meddlesome help. Ancient Chinese sound-phenomena, which are unrecognizable in the modern form of the loan-words, often peep out in the written form. Ancient Chinese *tieu* is now in Japanese *cō*, but is spelt *te-u*; Ach. *kieu* is *kiō*, but is spelt *ke-u*. Ach. *tieng* and *čiang* are both *čang* in Korean, but the former is spelt *tieng*, the latter *čiang*.

What makes Sino-Japanese especially valuable is that it exists in two versions. First the Japanese borrowed from S. E. and E. China, the so-called Go-on-version (5th and 6th centuries), then from N. China, the so-called Kan-on-version (7th and 8th centuries), and these two inter-independent versions are precious; the information afforded by one is often filled out in a very fortunate way by the data of the other. Thus, they show us that certain modern dialectal differences go back more than a thousand years. One example: Certain dialects in Shansi, one of the northern provinces, have *nd-*, *mb-* (*ndi*, *ndu*, etc.), answering to the *n-*, *m-* of the rest of China (*ni*, *nu*, etc.), which seems to clash with the rule about the absence of consonant groups in Chinese (see above p. 12), but only indicates an anticipatory closing of the nasal air passage in the transition from the nasal *n* to the oral *i n(d)i*. One would be tempted to think this to be quite a young phenomenon, were it not for the fact that Kan-on, which is based on Northern Chinese, regularly has *d-* for Chinese *n-* (Pek. *nan*, Kan-on *dan*, etc.) and *b-* for Chinese *m-* (Pek. *mi*, Kan-on *bi*, etc.), while Go-on, which is based on the old language of E. China, always has *n-* and *m-* (*nan*, *mi*, etc.). It is therefore proved through Sino-Ja-

panese, not only that this peculiarity *n* -> *nd*-, *m*-> *mb*- already existed in the 7th century, but actually that it was then, as it is now, confined to N. China (probably over a greater area than now).

When now, on the one hand, we have a sure point of departure in the old sources, in that we know which words were phonetically associated, which words formed phonological groups; and on the other hand we are enabled to take an old category of words like this, and examine, firstly, how it is reflected in Sino-Korean, Sino-Japanese, and Sino-Annamese, and secondly, how the words of this group appear in a great number of widely differing dialects; then it goes without saying that a linguist can with tolerable certainty reconstruct the sounds in these Ancient Chinese phonological groups. A beginning was made in the middle of the 19th century by an English missionary, Joseph Edkins, who propounded, together with many fantastic ideas, several correct and valuable theories. In 1898 the Dutchman Schaank made several important finds, and in the first decades of this century the French scholars Paul Pelliot and Henri Maspero drew up provisional systems for the interpretation of the Ancient Chinese sound-system, with many new and important ideas. The question was now taken up by me to its full extent, the dialects being systematically utilized for the first time, and after a further discussion between myself and Maspero, as a result of which I improved my system of interpretation in certain respects, thanks to his ideas, we have reached the point that the pronunciation of 6th century N. Chinese is now reconstructed but for a few details still to be discussed. I will give a few examples of the methods adopted for this reconstruction work.

We saw just now that the vertical columns in the rime-tables are in batches of four. If we study the words in the columns with the aid of the modern dialects, we shall have no difficulty in filling up three of them:

| *m* | | *p*<sup>•</sup> | *p* | *n* | | *f*<sup>•</sup> | *t* | *ng* | | *k*<sup>•</sup> | *k* |

Most of the dialects tally here, e. g.

	難	灘	單	岸	看	干
Peking	<i>nan</i>	<i>f an</i>	<i>tan</i>	( <i>an</i> )	<i>k an</i>	<i>kan</i>
Canton	<i>nān</i>	<i>f ān</i>	<i>tān</i>	<i>ngon</i>	<i>k on</i>	<i>kon</i>
Shanghai	<i>nā</i>	<i>f ā</i>	<i>tā</i>	<i>ngō</i>	<i>k ō</i>	<i>kō</i>

But the third column from the right is more interesting. In most of the Chinese dialects voiced initial plosives (*b*-, *d*-, *g*-) are lacking. An exception is formed by certain dialects in E. China, e. g. Shanghai, and in the Go-on-version of Sino-Japanese. If now we examine the Shanghai dialect, we shall find that the very words which in the rime-tables stand in the third column, have in Shanghai regularly *b*-, *d*-, *g*-, (or *dj*- < *g*-), whereas words which stand in columns 1 and 2 regularly have voiceless sounds, *k*-, *k*<sup>•</sup>-, *t*-, *t*<sup>•</sup>-, *p*-, *p*<sup>•</sup>-, exactly as elsewhere. And also in the old Sino-Japanese Go-on-version it is just the words of the third column of the rime-tables that have *b*-, *d*-, no others. Here we have examples (taken from different rimes):

	滿	件	判	半	難	壇	灘	單	擬	技	起	几
Shai	<i>me</i>	<i>be</i>	<i>p<sup>•</sup>e</i>	<i>pe</i>	<i>nā</i>	<i>dā</i>	<i>f ā</i>	<i>tā</i>	<i>nyi</i>	<i>dji</i>	<i>či</i>	<i>či</i>
Go	<i>man</i>	<i>ban</i>	<i>han</i>	<i>han</i>	<i>nan</i>	<i>dan</i>	<i>tan</i>	<i>tan</i>	( <i>gi</i> )	<i>gi</i>	<i>ki</i>	<i>ki</i>
Pek	<i>man</i>	<i>pan</i>	<i>p an</i>	<i>pan</i>	<i>nan</i>	<i>f an</i>	<i>f an</i>	<i>tan</i>	( <i>i</i> )	<i>či</i>	<i>či</i>	<i>či</i>

This makes it absolutely certain that all Ancient Chinese words in the third column of the rime-tables, and moreover all words spelt with these forms in the fan-ts'ie spellings, had voiced initial *b-*, *d-*, *g-*, etc.; and that when all these words appear in the majority of the dialects with voiceless initials, *p-*, *p'*-, *t-*, *t'*-, etc., this is the result of a sound-development away from the Ancient Chinese sound-system, which is therefore now only reflected in the said eastern dialects and in Go-on. As far back as a century ago, Marshman, a grammarian, guessed that the *b-*, *d-*, *g-* initials of Shanghai were remnants of an older phase of the language, but this was not proved until Edkins showed that this group of words with voiced plosives in Shanghai corresponded exactly to a category of words in the old sources, sharply differentiated from other categories, spelt with a special series of signs, and placed in the third column of the rime-tables. It should be added that I have shown the Ancient *b-*, *d-*, *g-* to have been aspirated *b'*-, *d'*-, *g'*-.

This was simple. But now we will take an example where all the given sources, fan-ts'ie and rime-tables, the foreign dialects, the modern dialects and old transcriptions, must be brought into play for a definite result to be arrived at.

Ancient Chinese was rich in *č* and *š* sounds. The rime-tables include, inter alia, 5 columns, which were already filled in by Edkins in this wise: *č-*, *č'*-, *š-(dj-)*, *š-*, *ž-* (we should correct *š-* to *ž'*-), which values were accepted by Pelliot and Maspero in their provisional systems. No pronouncement is here made about whether these sounds were hard, supradental («cerebrals»), *tš-*, *tš'*-, *š-*, etc., somewhat resembling English heartshaped; or whether

they were soft, mouillés (palatals), *tš-*, *tš<sup>c</sup>-*, *š-*, as in Italian *citta*.

When, with this as a starting-point, I set out to deal with the modern dialects, I found that in a number of otherwise sharply differentiated dialects each of the initials *č-*, *č<sup>c</sup>-*, *ǰ-*, *š-*, was treated in two different ways, and this quite independently of the following vowel. In other dialects this was not the case. Thus, for instance, in Kueihua (Kh), in most northerly Shansi, we have certain words that are read *tsäu*, others that are read *tšäu*, but they are under one and the same initial in the rime-tables, *č-*, and belong to one and the same ancient rime (A.Ch. *-i<sub>2</sub>u*). And in a dialect Hakkā (Hk), in the most southerly province of China, a dialect which has very little in common with the Kueihua, and also in Sino-Annamese (An., as pronounced in «Haut-Annam») we find exactly the same difference:

繡	A.Ch. <i>č<sub>2</sub>i<sub>2</sub>u</i> — Kh. <i>tsäu</i> — Hk. <i>tsiu</i> — An. <i>tš-</i>	} Pek. <i>tšou</i>
周	A.Ch. <i>č<sub>2</sub>i<sub>2</sub>u</i> — Kh. <i>tšäu</i> — Hk. <i>tšiu</i> — An. <i>t'-</i>	

This arouses our suspicions, and we turn to the fan-ts'ie spellings. There we find that within the initial *č* there are two sharply differentiated series of spellings, one in words that in Kueihua have become *tš*, in Hakka *ts-*, in An. *tš-*, the other in words that in Kueihua have become *tš-* in Hakka *tš-*, in An. *t'-*. To the distinction in these dialects corresponds therefore rigorously a distinction in the spellings of the 6th century, and thus we know that we have two Ancient Chinese *č-*, two *č<sup>c</sup>-*, two *ǰ-*, and two *š-*:

A.Ch. <i>č<sub>1</sub></i> ( <i>-i<sub>2</sub>u</i> etc.) — Kh. <i>ts-</i> — Hk. <i>ts-</i> — An. <i>tš-</i>	} Pek. <i>tš-</i> , Foochow <i>tš-</i>
A.Ch. <i>č<sub>2</sub></i> ( <i>-i<sub>2</sub>u</i> etc.) — Kh. <i>tš-</i> — Hk. <i>tš-</i> — An. <i>t'-</i>	

What was now the exact nature of these  $\check{c}_1$  and  $\check{c}_2$  in Ancient Chinese? That  $\check{c}_1$  should have been something in the style of ordinary dental  $ts-$  (as in Kueihua) is out of the question, for  $ts-$  also existed in Ancient Chinese, and stands in quite a different column. Was one hard ( $ts-$ ), and the other soft, mouillé ( $t\acute{s}-$ ), and in that case which was which? Peking has the hard sound  $ts-$  for both; Foochow has the soft sound  $t\acute{s}-$  for both. And  $\check{c}_2$  has the hard sound  $ts-$  in Kueihua, but the soft sounds  $t\acute{s}-$  in Hakka and  $t'-$  in Annam. It is evidently impossible to look to the dialects to decide the question.

But we are not at our wits' end. First we have the old sources, which go to prove that  $\check{c}_1$  was hard, supradental,  $t\check{s}$ , for in the rime-tables it stands in the same horizontal column as hard consonants,  $k-$ ,  $p-$ ,  $l-$ , etc., while  $\check{c}_2$  must have been soft, palatal,  $t\acute{s}-$ , for it stands side by side with palatal consonants,  $kj-$ ,  $pj-$ ,  $lj-$ , etc.<sup>1</sup> And the proof that this in itself somewhat risky analogical conclusion is correct, is afforded by the old transcriptions. It can be most clearly seen if, instead of the affricate  $\check{c}$  we take the fricative  $\check{s}$  as an example. For this we get the same doubling in the old fan-ts'ie spellings and in the dialects:

A.Ch. $\check{s}_1$ —	Kh. $s-$ —	Hk. $s-$ }	Pek. $\check{s}-$ , Foochow $s-$
A.Ch. $\check{s}_2$ —	Kh. $\check{s}-$ —	Hk. $\check{s}-$ }	

Since  $\check{s}_1$  (like  $\check{c}_1$  above) stands in the rime-tables side by side with  $k-$ ,  $l-$ , etc., and  $\check{s}_2$  (like  $\check{c}_2$  above) stands side by side with  $kj-$ ,  $lj-$ , etc., conclusion by analogy would make  $\check{s}_1$  hard, supradental  $\check{s}-$ , and  $\check{s}_2$  soft, palatal, mouillé  $\acute{s}-$ .

<sup>1</sup>) That these were hard in certain horizontal columns, and palatal in others, can be proved; this is not the place to adduce the proof; cf. my *Phonologie Chinoise*.

The ancient transcriptions confirm this most excellently. In the Buddhist transcriptions in the numerous translations of the holy books and other religious writings from about 400 to 900 A. D., we find consistently that:

Sanskrit *ṣ* is rendered by *ṣ*<sub>1</sub>, i. e. according to my interpretation *ṣ*-,

Sanskrit *ś* is rendered by *ś*<sub>2</sub>, i. e. according to my interpretation *ś*-

E. g.:

(*Gho*)*ṣa*, (*Pu*)*ṣya* is transcribed 沙 with *ṣ*<sub>1</sub>, i. e. according to me ACh. *ṣa*;

*śa*(*rīra*), (*vāi*)*śya* is transcribed 奢 or 舍 with *ś*<sub>2</sub>, i. e. according to me ACh. *śia*.

Examples to verify this occur by the dozen. And by this the division of the Ancient Chinese *č*-, *č'*-, *č''*-, *ś*- sounds into a supradental series *tṣ*-, *tṣ'*-, *dṣ'*-, *ṣ*-, and a palatal, mouillé series *tś*-, *tś'*-, *dś'*-, *ś*- is definitely proved.

The examples given may suffice to show the methods Sinology has to work with, when it is a question of reconstructing the language of the 6th century. It lies outside the scope of this work to go over the whole field and set out the total results of these investigations. But one more point is worth calling attention to. We have seen that we must work with a whole group of sources, the old rime-lexicons, the fan-ts'ie spellings and rime-tables, on the one hand, the modern dialects and the, so to say, foreign dialects in olden times, Sino-Korean, Sino-Japanese and Sino-Annamese, on the other. Now we must emphasize that these last-named dialects are often of extraordinary importance, since, provided you can see through and disregard the later phonetic developments on



the foreign soil, they reveal the state of the Chinese sounds at the time when these masses of loan-words were borrowed. It is not least this group of material that enables the investigator of to-day to attain sure results, which would have been out of the reach of the Chinese philologists of the 16th and 17th centuries, who had not access to these sources, or in any case did not understand their value.

As a matter of fact, it is often these foreign dialects that disclose facts which are not revealed by a single native Chinese dialect. We have, for instance, a large group of words with a certain final *-x*: *kx*, *tx*, *lx*, etc., some hundreds of words with this final *-x*. In the modern dialects they have very varying finals, though predominantly *-o*. If we examine the dialects of Peking (Pk) in the N. E., Tat'ung (Tt) in the N. of Shansi (N. China), Hinghien (Hh) and Wenshuei (Ws) in the centre of Shansi, Sianfu (Sa) in the N.W. (Province of Shensi), Nanking (Nk) in the centre, Shanghai (Sh) in the E., Swatow (Sw) in the S.E. (Prov. of Fukien), and Canton (Ca) in the S., we shall find:

				歌	多	羅	
Tt	Sa	Nk	Sw	Ca	<i>ko</i>	<i>to</i>	<i>lo</i>
				Pk	<i>kə</i>	<i>to</i>	<i>lo</i>
				Hh	<i>kə</i>	<i>tə</i>	<i>lə</i>
				Ws	<i>kbl</i>	<i>tbl</i>	<i>lbl</i> ( <i>bl</i> as in Russian)
				Sh	<i>ku</i>	<i>tu</i>	<i>lu</i>

There is, however, no dialect with any *-a* in these words. But if we go to the foreign dialects, it turns out that all three, though absolutely independent of each other, have *-a* in all these words:

# 歌 多 羅

Kan-on, Go-on	<i>ka</i>	<i>ta</i>	<i>ra</i>
Korean	<i>ka</i>	<i>ta</i>	<i>na</i>
Annam	<i>ka</i>	<i>ḍa</i>	<i>la</i>

There are therefore very strong reasons for considering these examples of how the old-time borrowers heard the words, to be conclusive, and to infer that all these words with the rime *-x* ended in Ancient Chinese in *-a* (for reasons I cannot here go into, it must have been an *ā* «grave» as in modern English *father*), *kā*, *tā*, *lā*, etc., that this *-ā* throughout China became *-o* (just as Old Swedish *gās* became *gōs*, written *gås*), and that this *-o* was further shifted to *-u* in Shanghai, *-u* in Wenshuei, *-ə* in Hinghien, and to *-ə* in Peking only after gutturals (*kə*, but *tə*). For this conclusion we have an excellent touchstone in old transcriptions, especially the innumerable Buddhist ones. There we find that this class of words in *-x* was regularly used to render the foreign *-a*; e. g. 摩訶摩耶 Pek. *mo-ḥə-mo-īā*, which according to my system of reconstruction was an Ancient Chinese *muā-ḥā-muā-īa*. It corresponds to Sanskrit *Mahāmāya*, and with that the question of this *-ā* rime is definitely clear.

What I would especially call linguists' attention to, is that the reconstructions of Ancient Chinese which we can achieve with these means, are not directly comparable with the reconstructions which, for instance, the scholars in the field of comparative Indo-European philology make of the Indo-European parent language. These are, in reality, merely a series of formulae, the expressions of certain correspondences between actually recorded languages, but formulae that do not aim at a certain fixed

point in time, and the parent language may have possessed a whole number of things that do not appear from the derived languages, and therefore do not find expression in these formulae. As Holger Pedersen<sup>1</sup> has expressed it: if merely from the forms of the modern Teutonic languages we were to try to reconstruct an Ancient Teutonic parent form for the word corresponding to the Latin *cornu*, i. e. from Eng. Germ. Scand. *horn*, Du. *hoorn* we should get *horn*, but certainly not *horna*, which we now know it to have been, for instance through the inscription on the famous golden horn from Gallehus, and otherwise. Likewise there may have been in Indo-European various things that none of the derived languages reveals, and which we therefore never reach by our reconstructions. Our Ancient Chinese reconstructions are of quite a different, and more real, concrete kind. We have in the rime-lexicons, rime-tables, and fan-ts'ie, a codification of the phonological groups of the language at a definite time; we have furthermore foreign dialects which reveal how the foreigners heard the words at about the same period; we have transcriptions of foreign words by Chinese characters made on the basis of the Chinese pronunciation during this very epoch; and it is the role of dialect investigation to fill out and confirm the conclusions respecting the concrete sound-system of the 6th century which all these sources enable us to set up. If we have rightly made use of these various, inter-supporting sources, we have got from them, not formulae of the abstract kind that the comparatists of Indo-European set up, and which do not aim at reflecting a concrete language at a certain given

<sup>1</sup>) H. Pedersen Sprogvidenskaben i det nittonde aarhundrede, 1924, p. 247.

point of time, but a faithful picture of how the author of Ts'ie-yün pronounced his language about 600 A. D. And there is no risk of any gaps, elements that have remained undiscovered. We get, for instance, reconstructive formulae like this: *kj<sub>i</sub><sup>w</sup>än*. This means that we have fixed the Ancient Chinese form of the word from the very beginning to the end, and we have determined 5 elements independent of each other: 1) the initial *k-* in contradistinction to *p-*, *t-*, etc., 2) the rime *-än* as opposed to *-ân*, *-an*, *-ən*, etc., 3) *j* («yod») as against forms like *kan*, *kien*, *kuən*, without *j*, 4) *-i-* as against forms like *kân*, *kuân*, *kien*, which had no subordinate *-i-*, and 5) *w* as against forms like *kân*, *kien*, *kjiän*, without *w*.

We have consequently succeeded in reconstructing in detail Ancient Chinese as it was spoken in the 6th century. Or, more rightly, a branch of the then spoken language; for dialectal differences existed even at that time. The dialect thus reconstructed was, however, one of the greatest and most important — it was the dialect of the old capital, Sianfu in Shensi — for it turns out that the great majority of present dialects, spread as they now are over an enormous area, can be organically explained directly from it, and descended directly from it (while spreading it must have superseded the other dialects, and formed a kind of *koinä* language). An exception is only formed by a few dialects in the province of Fukien, which in certain cases (like the Sino-Japanese version Go-on) point to an early dialect deviating from the language of Ts'ie-yün.

By the reconstruction of the Ts'ie-yün system much has been won, especially for the historians; for it was precisely during the first thousand years of our era that the Chinese were in very close communication with the peoples

and cultures of the rest of Asia; and this knowledge of the pronunciation of the 6th century is a fixed point of great value in interpreting all sorts of transcriptions of foreign words and names. But this result is, and must be, but a beginning. Chinese literature extends far back into pre-Christian times, and from about 500 B. C. onwards it is abundant and valuable. An inevitable requirement is, as I have already pointed out, to come to close quarters also with these earlier stages of the language, and if possible reveal it as it was spoken, despite the obstacles that the curious script sets in our path.

It is the more important to submit the oldest texts to all the linguistic analysis that is feasible, and this because it is a hitherto untested, but, as we shall see later, an assuredly possible way of testing their genuineness. There is the question whether the Chinese classics, the oldest literature down to 200 years before Christ, is to an essential degree faithfully preserved, so that what we now have access to is, with minor inevitable modifications, a body of genuine documents from that time, or if and to what extent it is a rehash of older fragments pieced together into books, or simply free forgeries by the learned, from the centuries immediately before and after Christ; and this question is as a matter of fact one of the most burning and difficult for Sinologists. It is so complicated because the fortunes of literature in China have always been anything but peaceful. I will now briefly describe these varying fortunes, and the difficulties they place in the investigator's way.

## CHAPTER V.

**I**t has already been stated that the earliest documents we possess, apart from a few bronzes, which may possibly be older, are inscribed pieces of bone and tortoise-shell from the end of the Shang-Yin dynasty (approximately 1766—1122). On these we already find a good clerical technique, and it is certain that longer documents were produced as early as 1000 B. C.: an imperial prince, Chou Kung (c. 1100), offered the gods his life as a votive offering, wrote down his prayer, the wording of which exists in an old document, and enclosed it in a stone chest, where important state documents were preserved. One of China's greatest thinkers, Lao-tsi, was according to tradition a keeper of the archives of the Chou dynasty in the 6th century B. C., and his younger contemporary Confucius, made anthologies of existing collections of folk songs and sacrificial odes, historical documents, etc., some of which seem to have been indited a chiliad and a half of years before Confucius.

This is of no imposing antiquity. From the near Orient we have documents preserved from much older epochs. We might therefore expect that many Chinese MSS from pre-Christian times should be preserved. Unfortunately,

this is not so, the reason, among other causes, being an historical event that the friends of literature in China have never ceased to lament. At the close of the feudal period in China, in the 3rd century B. C., when the country became solidified into a great centralized power of a more modern constitution, the originator of this unification, the Emperor Ts'in Shi-huang-ti, found his bitterest opponents among the literati, who attacked him with their fierce criticism and their panegyrics on the good old times as they appeared in literature. The Emperor solved the difficulty by decreeing a general burning of books in 213 B. C., with severe punishment for those who hid books; and this seems to have been most effective. Apart from what was preserved in the Palace library, the stock of books in China seems to have been almost entirely destroyed. In the struggles that soon afterwards led to the fall of the Ts'in dynasty, and the setting up of the House of Han in 206 B. C., the Palace was devastated, and with that the archive copies were also lost. During the following centuries all was done that was feasible to search out what had escaped destruction, and the movement had invaluable help in the mode of study that had always been in vogue in China, viz. the committing to memory of important documents. The finds that were made were verified and interpreted, thanks to the oral traditions of the various schools of learning, traditions that were passed on from teacher to pupil generation after generation. Thus was written down a considerable body of literature, which is ascribed by the learned world in China to the epoch before the book-burning, especially the Confucian classics, i. e. historical documents and poetic productions collected by Confucius, Confucius' conversations as taken down by his

3  
pupils' pupils, the master's first great follower's writings (those of Mencius, from about 300 B. C.), several other great historical works from approximately the same time, several philosophical works not connected with the Confucian school, and certain comprehensive notes on rites and ceremonies.

This salvage work gave rise to philological studies in China. Men were trained as critics of texts, exegetists, palæographers; and especially in the 2nd century A. D. there arose a school of distinguished teachers who were to exercise a decisive influence on scholarship in the subsequent history of China.

In handing down the old writings, however, the purely technical development came to be of great, and as we shall see, of disastrous importance.

From having originally engraved with sharp instruments on bone and tortoise-shell, the scribes had advanced, a good long while even before the book-burning, to writing with a wooden pen in a kind of ink upon wood, bamboo, or occasionally silk; and soon after the book-burning the hair-pencil was invented, and has remained to our own times the writing-implement of the Chinese. But wooden documents were clumsy and bulky, silk was dear, and so they began to experiment in producing cheaper material with the same good qualities as silk, until at last in 105 A. D. the clever Ts'ai Lun invented the cheap paper. Long before Ts'ai-Lun, paper made out of waste-silk had already been produced. It was soaked and beaten into a pulp, a thin layer of which was caught on a kind of thin mat, and was dried into a kind of silk-paper. However, even this material was too dear and difficult to obtain, and Ts'ai Lun's great discovery was that cheap material could be



used instead of the silk substance, and that better results accrued; this was the bark of the paper-mulberry, hemp buds, rags, and old fish-nets. Of these he prepared a pulp which he treated in the above-mentioned way, in thin layers, and thus produced a supple, even, thin writing-material of the cheapest kind. His invention met with the warm approval of the Emperor, and spread rapidly. In the ancient little oasis town of Lou-lan on the Lopnor, Sven Hedin, who discovered the place and made excavations there, collected a number of specimens of early paper, and some of these fragments go back to the first century after Ts'ai Lun's invention. The same holds good of certain finds made by Aurel Stein at the Great Wall in the extreme N. W. of China. That such documents should have found their way to distant Lou-lan, where thanks to the desert drought they escaped the fate that overtook paper documents of that period in China proper, is explained by the circumstance that a caravan route between China and the near East, much used especially for the silk trade with the Roman Orient, passed through Lou-lan.

This invention of paper at such an early age, was a great misfortune for China. This may sound paradoxical, but is nevertheless a fact. China, as we know, was the first country to invent, among so many other things, the printer's art, but this was comparatively late, and regular printing on any big scale did not begin until the 10th century A. D.

The circumstances were therefore quite different from those in Europe. Here we began to make a serious use of paper at quite a late date. The art of manufacturing it had reached the near East from China in the 8th century, and from there the novelty spread very slowly westwards,

not coming into general use in Europe before the 13th century. Before that one had regularly written on parchment. But in Europe the art of printing was already practised in the 15th century, consequently very soon after the introduction of paper; and this has been of extraordinary importance to western literature. The quickly multiplying art was obtained soon after the appearance of the fragile material, paper. There was therefore no breach in continuity: we have documents preserved from every stage. Following a period of MSS which were few, but imperishable owing to the strong material, parchment, there came, practically without an interval, a period of book-production with weak material, but kept from obliteration by the great numbers of copies.

These favourable conditions did not exist in China. We have seen that the oldest stock of wooden books was almost obliterated by the book-burning of 213 B. C. The new copies on wood that were made in the next few centuries cannot have been excessively numerous, since they had to compete with documents on silk. Of this new stock of wood books there is therefore nothing left, with the exception of a number of fragments rescued by Stein and Sven Hedin from their hiding-places in Central Asia. When the really broad learning blossomed out in China from the 2nd century A. D. onwards, Ts'ai Lun's invention of paper had already spread everywhere, and even if they continued to write on wood by the side of paper, it can never have been a question of many copies.

But from the invention of the weak material in 105 A. D. to the serious adoption of printing in the 10th century, there is a dreadful gap of many centuries. During this long period they copied the old texts by hand, thus in a

necessarily small number of copies. But at the same time the material of these copies had no resisting power, like western parchment, but was most perishable; and the main body of the written copies of Chinese literary productions that were made on paper during this long period of time, has disappeared without a trace, owing to the comparative paucity of the copies and their fragility. The same is of course true of the copies written on silk after the book-burning episode. It is therefore as natural as it is sad, that from all the period that precedes the 10th century, the great bulk of MSS has been hopelessly lost. An exception is formed by the finds made by Hedin and Stein in Chinese Turkestan, documents on wood and paper; and another is formed by a small number of MSS on paper and silk from the first millennium A. D., which, despite all vicissitudes, have succeeded in coming down to our own day. An important part of these last treasures consists of a library which has been preserved since 1035 A. D. in a walled-up vault of a temple in Tun-huang, in the most north-westerly province of China, Kansu — a precious store of books, the most important of which have been acquired by Stein and Pelliot, and transferred to London and Paris. Of course there are some privately owned MSS in China, and some to be found in Japan. But all this is an infinitely small fraction of the huge store of books in the shape of paper MSS that existed and was lost between Ts'ai Lun in 105 A. D. and the printing activity of the 10th century.

Fresh transcripts have preserved from generation to generation the most important works, such as the classics, the dynastic histories, some lexicons and encyclopædias, etc.; masses of other interesting work have disappeared.

Fragments of them are occasionally to be found in the shape of quotations in later works, made at a time when copies of the publications were still extant. And from such quotations we can often see that many a work which would have been invaluable to research, is unfortunately lost for ever. The Chinese scholars of a later age have been at great pains to piece together quotations of this kind, and thus reconstruct lost, valuable works.

If we add to this, as the last stones in the structure, the facts that the technique of the first few centuries of printing was naturally primitive, and editions limited; and that during the fearful wars that were waged, first when the Jinghiskhanides conquered China in the 13th century, and then when the Manchus ravaged, burnt and subjugated the country in the 17th century, mighty treasures of these early impressions, the incunabula, so to speak, of China, were lost; we can understand that Chinese literature shows lamentable gaps.

It stands to reason that during the long and fluctuating career of Chinese literature, the handing down of the old texts presents many obscure points, and that a priori we must take up a somewhat sceptical attitude towards them in that form of theirs which became crystallised as the traditional one during the last thousand years. To take a concrete example, I will touch on the chief points in the fortunes sustained by China's oldest preserved work, *Shu-king*, a collection of historical documents, now that those fortunes have been brilliantly investigated by Paul Pelliot, who followed up the preliminary researches made by the Chinese textual critics and Ed. Chavannes.<sup>1</sup>

<sup>1</sup>) Paul Pelliot, *Le Chou king en caractères anciens et le Chang chou che wen; Mémoires concernant l'Asie Orientale II.* Paris, 1916.

The generally accepted tradition as to the handing down of Shu-king, as established by two authoritative scholars in the 7th century A. D., K'ung Ying-ta and Lu Tê-ming, is as follows. The great Confucius, who flourished about 500 B. C., collected historical documents into a Shu-king of 100 chapters. Destroyed in the book-burning of 213, the book was partly rewritten in the reign of Hiao-wên (179—157), an old scholar, Fu Shêng, reciting from memory 29 chapters, which were written down in the then current simplified script, called *kin-wên*, i. e. 'modern writing'. About 100 B. C. a descendant of Confucius, called K'ung An-kuo, obtained possession of an copy of Shu-king, written in old-fashioned characters, the so-called 'tadpole characters', a copy that was found in a wall in pulling down Confucius's house. He compared this with Fu Shêng's text, and found that the 29 chapters of the latter were actually, owing to faulty partitioning, 34 of Confucius's 100 chapters, and in the newly discovered copy there were in addition 24 other chapters, and one more that was a kind of introductory chapter. K'ung An-kuo now wrote a preface and added a commentary to the whole. His work got the name Shu-king in *ku-wên*, i. e. in ancient writing. Fu Shêng, as well as K'ung An-kuo, got pupils and pupils' pupils, who propagated their versions of Shu-king. During the political chaos in the 3rd century A. D., traces of the handing-down are lost, but about 320 A. D. a learned man, Mei Tsê, presented the Emperor with a Shu-king that was supposed to be K'ung An-kuo's version, and it is this edition of Mei Tsê's that became the traditional text, commentated by the two aforesaid scholars in the 7th century, and still to-day widely regarded as authoritative, and studied by every student.

This tradition about the preservation of Shu-king, rests entirely on the preface in Mei Tsê's version, the genuineness and truth of which the said scholars did not question. But already from the 12th century onwards there arose a body of learned men with a more critical eye, who showed that the alleged K'ung An-kuo preface and commentary could not have been written in K'ung An-kuo's time, and that Mei Tsê's copy was a forgery made at the end of the 3rd or beginning of the 4th century. The real K'ung An-kuo's edition has disappeared without leaving a trace. The pseudo K'ung An-kuo had taken Fu Shêng's 29 chapters, rearranged them somewhat, and added various fragments from other sources, which he stewed into a number of chapters that he ascribed to Confucius's collection, and called it K'ung An-kuo's edition. As a matter of fact, if we go to the great historian Sî-ma Ts'ien, who was a contemporary of K'ung An-kuo, we shall find that he mentions Fu Shêng and his preservation of the Shu-king text, but he says expressly that that nonagenarian, who when the books were burnt had hidden his Shu-king, and afterwards only recovered 29 chapters, confined himself to teaching these 29 chapters, and consequently did not rescue Shu-king out of his own memory. The same historian mentions K'ung An-kuo's version of Shu-king as a precious possession of Confucius's family, but he has nothing at all about the story of the find being made when Confucius's house was pulled down. Pelliot shows that these embellishments and additions to the story of Shu-king's rescue originated in the 1st century B. C., and were culled from this time onwards from many sources and woven into the pseudo K'ung An-kuo's preface. And neither the historian Sî-ma Ts'ien, nor other sources during the next few

centuries state that the real K'ung An-kuo wrote a preface or a commentary. He only deciphered the manuscript belonging to his family, and preface as well as commentary are exploits of the falsifier.

Now when this alleged K'ung An-kuo edition, supplanting the honest Fu Shêng edition (on which it was partly based, but with minor alterations) became more and more a standard work, the old, but as I have said, false tradition that whereas Fu Shêng had written down from memory his 29 chapters in *kin-wên* 'modern writing', K'ung An-kuo's version was in *ku-wên* 'ancient writing' came to have disastrous effects. It was fatal in this wise, that in the centuries immediately following Mei Tsê's presenting of the false document, various scholars in copying and teaching the text smuggled in unusual, quaint variants for the normal characters in the text; this, in order to justify the tradition and the name 'Shu-king in ancient writing'. In this manner, not only were the last parts tampered with, but also the essentially authentic ones based on Fu Shêng. This point had been reached when, in the 7th century, Lu Tê-ming and K'ung Ying-ta took this roughly handled and in part false text as the basis for their learned and for all time authoritative commentaries.

But this is not the end of the tragicomic story. As far back as from 400 A. D. onwards, individual scholars made certain attempts to get rid of these archaizing characters, either because they suspected them of having been falsely smuggled in, or because, for practical reasons, they wanted to substitute for them current characters. The matter was settled by the Emperor Hsuan-tsung of the T'ang dynasty, who thought the variations between the versions of different scribes had become alarming; he resolved to

normalize the text. At his command a learned man, Wei Pao, set to work, and in 744 A. D. produced a Shu-king entirely transcribed in current script and cleansed of all archaizing characters. This normal edition gained a complete success, and the earlier ones soon disappeared from circulation.

For the lover of literature this was, however, a regrettable matter. After all, the unrevised edition with the archaic characters was older and more original, and a few learned men upheld the tradition by bringing out new editions of it in the next few centuries. But these were isolated phenomena, and from the 14th century onwards all the copies of Mei Tsê's edition (i. e. the pseudo K'ung An-kuo from circa 300, with later substitutions of archaic variants) disappeared. The only remaining trace was a commentary to an edition of 1082 written in the 12th century; and with the help of this commentary, scholars of later times have tried to reconstruct the archaizing Mei Tsê text that goes back to earlier than 744. In Japan, too, they thought they had, in documents preserved there — an edition of 1322 and a couple of MSS from the 14th century that are based on older publications — reached back to the time before the great normalising of 744. However, when a lucky chance ordained that Stein and Pelliot should bring back from the temple in Tun-huang a number of MSS from the T'ang epoch (618—907), there were among these several fragments of Shu-king in the version prior to the normalising of 744. These have not yet all been published and worked up, but they show that not only the reconstruction on the basis of the commentary of the 12th century, but also the Japanese versions are unreliable; they contain far too great a percentage of archaizing characters, compared with the genuine T'ang documents from Tun-huang.



To summarise the results:

Of Confucius's original collection of documents, such as it had been copied and studied circa 500—213 B. C., we know of only two copies that survived the book-burning: one belonging to Fu Shêng, in 29 chapters; the other belonging to K'ung An-kuo. The former MS is lost, and the school that from 200 B. C. to 200 A. D. approximately took its stand on the basis of this version, died out. The K'ung An-kuo version, which his contemporary Si-ma Ts'ien had actually seen, has disappeared without leaving a trace. Circa 300 A. D. appears a pseudo K'ung An-kuo version, based on Fu Shêng's version (with minor alterations to make it look like an independent version), and with spurious additions. During the centuries that followed, a lot of archaizing characters were introduced into this already corrupt text. And owing to the normalising of 744, we do not know for certain how even this modified Fu Shêng version, tampered with by the falsifier and decked out with archaizing characters, looked, except for the fragments found at Tun-huang.

The question is ready to hand: How can we work with such material in our investigations about Old China? What guarantees have we that anything at all is reliable in texts which have suffered so badly in the handing-down process?

But we must admit, to begin with, that the above is an extreme case. Most of the other problems are not nearly so complicated. As regards a number of texts, we can say with certainty that they are spurious; as regards others, we can say with tolerable certainty that they are genuine. Yet there are a number of works, some of which belong to the most important of early Chinese literature, where we

have not been able to decide whether they are spurious or genuine. Here, it seems to me, linguistics should lend a helping hand, and try to contribute to the attainment of sure results.

---

If, then, we are not satisfied with the results already obtained, the reconstruction of a Chinese chief dialect in the 6th century A. D., but wish to go further back and treat, by linguistic methods, documents that are much older yet, we meet with difficulties which, in the light of what I have mentioned about the Chinese language and script, appear unsurmountable. There we have before us texts like a picture-gallery, speaking to the eye but not to the ear. The writing does not inform us how the authors pronounced their words, and we cannot therefore examine words, stems and roots, as in other languages. If there were dialect differences, different pronunciations of the same word in different texts, this is carefully hidden behind the blank walls of the Chinese ideographs. Endings and prefixes can hardly have existed to any considerable extent; if they once existed, as is probable, they had already been lopped away to an essential degree. The accident therefore gives us no help. Not even the choice of words gives us a sure starting-point for dating a text, for literary Chinese, as has been pointed out above, is extremely syncretic: a word continues to live and be used for centuries, even when it no longer belongs to the spoken language; the vocabulary is therefore immense, and without definite limits to its periods. Where is the unfortunate linguist to search in order that he may find cultivable ground that will yield a crop?

The situation is difficult, but it is not hopeless. Even

this literature, apparently so well armoured against the attacks of all linguistic investigations, has its vulnerable points. Hitherto they have scarcely been searched for, but I will give an example of one or two possibilities.

Confucius made an extract from the court annals of his native state, the principality of Lu, under the title of *Ch'un-ts'iu* 'Spring and Autumn'. This is a short, excessively concise collection of documents of very slight literary importance. It is provided with a full commentary, *Tso-chuan*, which for every entry in the dry chronicle gives a broad and full-flavoured historical sketch, which gives us a full Chinese history that extends from 722 to 468 B. C. Round this *Tso-chuan* the waves of strife have run high. According to the traditional theory, it was written by a contemporary of Confucius's, Tso K'iu-ming, a person of whom we otherwise know practically nothing. *Tso-chuan* is supposed to mean Tso's commentary. As, however, the word *tso* means 'left', the possibility remains that *Tso-chuan* should be translated as 'commentary to the left', and the learned German Sinologist, W. Grube, has suggested that *Tso-chuan* must have been written by Confucius himself, with the *Ch'un-ts'iu* annals merely as a theme, which would be a happy explanation of why Confucius himself speaks of his *Ch'un-ts'iu* as the chief work of his life; this would be difficult to explain, if, by *Ch'un-ts'iu*, he had only meant his dry extracts of annals. This theory, tentatively thrown out by Grube, has been confuted with strong arguments by Otto Franke, who explains Confucius's view of his *Ch'un-ts'iu* in quite another way<sup>1</sup>, and regards

<sup>1</sup>) The *Ch'un-ts'iu* text was the written theme of an esoteric doctrine, handed down orally, which was afterwards recorded chiefly in certain other classical writings, the commentaries of Kung-Yang and Ku-liang, and the *Ch'un-ts'iu fan lu* by Tung Chung-shu.

Tso-chuan as a later, independent, historical work, which long afterwards was cut up and fitted together with Confucius's Ch'un-ts'iu. His reasons, however, have not convinced certain younger German Sinologists, who embrace and defend Grube's theory about Confucius's authorship with greater enthusiasm than the suggester himself. And what is more, the facts about the preservation of Tso-chuan after the book-burning are obscure, and some radical critics consider it to be a forgery. In 1912 a Japanese scholar tried by astronomical proofs to show that it was put together during the Ts'ien-Han dynasty (206 B. C.—24 A. D.), and a school of modern Chinese critics, with K'ang Yu-wei at the head of them, also consider it to be a pure forgery from the 1st century B. C. Otto Franke does not, it is true, go so far as this, but he considers the dismembered text to have been seriously tampered with.

In this much debated text there is one linguistic phenomenon that we will now examine.

In classical Chinese there are two words 於 and 于, both now pronounced *yü* in Mandarin, which serve as localizing prepositions, meaning 'in, on, at, into, with (auprès de)'; and then in an extended use 'vis-à-vis, with regard to, in relation to', etc: *yü king* 'in the capital', *yü ch'ao* 'in the court, at the court', *yü shī* 'in the market', *yü shan ting* 'on the top of the mountain', *yü wang* 'with (auprès de) the king', etc. Of the two chief European grammarians of classical Chinese, Stanislas Julien (1869) and G. v. d. Gabelentz (1881), the latter declares them to be «identical», i. e. clerical variants of the same word; the former is more prudent, and says they are synonymous. They are certainly not identical. As late as the 6th century A. D. (a stage in the language that we have been able

to reconstruct; see above), 於 was *i<sup>wo</sup>*, while 于 was *jiu*, and this *jiu* had sprung from an older *gjiu*, as we know through sources I shall return to later. But Julien is perfectly right in saying that they are synonymous. If we take the Confucian classics *in their entirety*, we shall find that we can get numerous examples of all these meanings, 'in, on, at, with', etc., for both 於 and 于. But if we take the texts *separately*, and examine them on this point — a very natural idea that does not seem to have occurred to anyone before — we get interesting results. In Shu-king, a text the vicissitudes of which were related in the last chapter, there is a practically exclusive use of 于. In the conversations of Confucius and his disciples, as also in the work of the philosopher Mencius (died 289 B. C.), 於 figures almost exclusively; the sporadic 于 are very few. This is already some light on the curious situation with two synonymous prepositions, viz. that one predominates in *one* dialect, the other in another. But if we take the above-mentioned Tso-chuan, the case is complicated in a most interesting manner. Here both occur, and at first sight perfectly *promiscuously*. Thus we find,

*Ts'ing shi* 于 *Chu* 'he demanded an army in (from) Chu.'

*Sung kung i pi ts'ing* 於 *Wei* 'The prince of Sung with presents made a demand in (from) Wei.' Cases of apparently complete synonymousness occur on almost every page in Tso-chuan, and it is therefore not to be wondered at that no one should have hit upon the idea of pushing the matter.

The question should, however, be put in a particular way. I have made statistics of the *unambiguous* cases where these

prepositions have been used in the following three meanings:

A. 'with (chez, auprès de, vis-à-vis)', mostly followed by the name of a person, e. g. *k'iu ch'ung yü chu-hou* 'he curried favour with the princes'; *kung wen yü Chung-chung* 'the duke asked *auprès de* Chung-chung.'

B. 'at (à)', with a place-name, e. g. *Pai Sung shī yü ch'eng Ying* 'he defeated the army of Sung at Ying';

C. 'inside, in, on, into (dans)', e. g. *yü tsung miao* 'in the ancestral temple', *chī yü kiao* 'he arrived in the suburbs';

I then found that while in the meaning C 於 and 于 are absolutely promiscuously used, almost equally frequently, 於 is the regular preposition in the sense A (about 7 於 to 1 于 in this sense); and 于 is the normal preposition in the sense B (about 5 于 to 1 於). So we see that the Tso-chuan dialect has commenced the process of mixing up 於 and 于, but that this process had not gone so far that we cannot clearly discern an interesting functional difference between them<sup>1</sup>. The distinction is not absolutely rigorous, but the fundamental truth of the rule is undeniable.

In the language of later times, when it was not suspected that such a distinction had ever existed, 於 *i<sup>uo</sup>* and 于 *jiu* were used quite *promiscuously*, as pure synonyms.

<sup>1</sup>) From the statistics are excluded all cases where *yü* is followed by the name of a *p r i n c i p a l i t y*, because these cases are obviously ambiguous. A phrase like *ts'ing shī yü Chu* can evidently be read in two senses. To put it in French: Il demanda une armée *dans* (le pays de) Chu (于, sense B); il demande une armée *auprès de* (la principauté de) Chu (於, sense A).

But in certain dialects the levelling began very early. While the distinction, as we have seen, is still quite noticeable in Tso-chuan, it has disappeared not only in the very old dialect on which Shu-king is based, but also in the language of the principality of Lu, the native soil of Confucius and Mencius. In the dialect of Shu-king 于 conquered, in Lu 於.

The yü prepositions are, after all, but a single detail in the grammatical system; one must not, of course, draw far-reaching conclusions from an isolated phenomenon like this. As a matter of fact it is also but one of a whole series of grammatical facts upon which I base my opinion (for details, see my paper «On the Authenticity and Nature of the Tso-chuan», Göteborgs Högskolas Årsskrift, 1926), facts that allow us to draw the most important conclusions:

(1) We get it confirmed that in archaic Chinese, a few centuries B. C., there were definite dialectical differences, a fact one had every reason to surmise, but the proving of which was made difficult by the script, which does not reveal whether a certain word was pronounced differently in different parts of the country. All we knew previously about the dialects from this old time, were some notes from the Han period on a number of special 'dialect words' used in certain regions. The dialect investigation here goes right into the heart of the classical literature.

(2) Confucius cannot be the father of Tso-chuan. It is of special importance that the differences between Confucius and Tso-chuan should bear upon *the use of frequently occurring grammatical auxiliaries*: in such intimate grammatical things, of which the speaker or writer was certainly unconscious himself, it goes without saying that a speaker

or author has not different principles on different occasions. What is more; the 2000-year old, ever authoritative tradition in China, that Tso-chuan was written by Tso K'iu-ming, *a man from Lu*, gets a shot below the water-line, for Tso-chuan is evidently not written in the Lu dialect.

(3) Most important of all, however, is this: whatever hazardous fortunes the Tso-*chan* text has undergone, it must in the main be genuine — not indeed in its script, but in its *wording* (though of course it may contain a number of interpolations). For a falsifier who wrote at a time when

於 and 于 had already been completely levelled functionally, and were perfect synonyms, *could* not have invented and carried through this grammatical distinction, nor would he have understood it and been able to imitate it, if it happened to exist in some other old documents. And the grammatical system, the *constellation of auxiliaries* in the Tso-chuan is quite unique, as I have shown in detail elsewhere (see my article just referred to). This is the best proof of its authenticity.

The treatment of 於 and 于 in Tso-chuan actually throws a vivid light on how the fortunate preservation of a text like this took place, in spite of a long series of copying by different scribes, and with one revised system of script after the other. In old times, and right into our own, scholars learnt the old texts by heart, handing them down from teacher to pupil through generations, and rewriting them in the successive *script forms*, but never changing *the wording*. The oral tradition has been the chief pillar, and the pronunciation of the text has gradually changed at the same pace that the spoken language changed — the ideographic writing did not, as we know, interfere



with this (see above, p. 58). But where a distinction existed which had not been lost through the levelling of sounds, we have the old difference still left in later times. Thus, for instance, the learned men of the 6th century A. D. still made a distinction, as a reflection of an old difference, between the pronunciation of *·i<sup>wo</sup>* and *j<sub>i</sub>u* for two prepositions that to them were synonymous, and with a remarkable fidelity to the old oral tradition they have — at any rate to a large extent — preserved through the centuries the *distribution* of 於 (*·i<sup>wo</sup>*) and 于 (*j<sub>i</sub>u*) in the texts.

And just this, that it is a question of such a trifle, such an insignificant and intimate grammatical difference in function, remains the most valuable thing of all: in more striking things one might imagine the tradition to have been kept, and the details to have been carelessly dealt with. Here, on the contrary, we see that the fidelity of tradition prevailed even with regard to the exact wording of the text.

If, therefore, in the matter of purely grammatical words, we can make quite productive investigations, even if we do not know how the language was pronounced, it becomes considerably more troublesome when we attempt to reconstruct, from the language of the 6th century, which we have now reconstructed, the sound-system of a thousand years earlier. But we are not entirely without expedients: the Chinese script is the great hindrance, but nevertheless it is in one respect a valuable assistant. We must remember that the great mass of the Chinese characters are composed of a significant and a phonetic. It easily suggests itself that we might, by studying *the choice of the phonetics in compound characters*, get some useful suggestions. It is,

of course, true that the Chinese characters were composed in the course of many centuries, but the method of making compounds of significant and phonetics did not begin on a large scale until comparatively late, as we can see from the old bronze inscriptions; and if we say that the chief mass of really current characters of this kind came into being during a period covering approximately 500—100 B. C., with an intensive period about 200 B. C., we shall be on the safe side. These signs should therefore be able to reveal something about the state of the sounds of that period. In a lexical work, «Analytic Dictionary of Chinese and Sino-Japanese», published by me in 1923, I was able to draw some far-reaching conclusions from this material, and I will here give some examples of how I set to work.

If, starting from the language of the 6th century, as reconstructed by us, we examine the series of compound characters formed with a certain phonetic, e. g.

干 *kân* phonetic in 肝 *kân*, 刊 *k'ân*, 旱 *yân*, 罕 *xân*; or:

白 *b'vk* phonetic in 百 *pvk*, 珀 *p'vk*, 泊 *b'ák*, 碧 *piäk*;

we shall find that there is conformity between the simple sign and its compounds, both with respect to initial sound and final sound. 干 *kân* has a guttural at the beginning and *n* at the end, and so have the composed characters.

白 *b'vk* has a labial at the beginning and a guttural at the end, and this is also true of 百 *pvk*, etc.

Moreover, we find for the most part cognate vowels. This general law holds good in 9 cases out of 10, even in the 6th century, and held good, with no shadow of doubt, when the signs were coined. For one character to serve as a sound-in-

dicator, or phonetic, in another, this measure of conformity in the pronunciation of the syllables was evidently demanded. When, then, in the series of phonetic compounds as they were pronounced in the 6th century, we find certain flagrant deviations from the main rule, e. g.: 甬 *i<sup>w</sup>ong* phonetic in 通 *t'ung* or 白 *b'vk* phonetic in 怕 *p'a*, we cannot refrain from asking whether anyone could be so modest in his demands on phonetic similarity as to let a *i<sup>w</sup>ong* be the phonetic in a *t'ung*, or a *b'vk* in a *p'a*. In the former case there is uniformity in the final consonants, but not in the initial sound; in the latter, only in the initial consonants (both labials), but not in the final. Would it not be more reasonable to assume that the striking difference in sound between 甬 and 通, and between 白 and 怕 is due to changes that have taken place in the language, and that they were more closely resembling in pre-Christian times, when the composed characters 通 and 怕 were created? We shall get an answer to the question without any great difficulty, if we examine a few series of words put together with one and the same phonetic. If 甬 *i<sup>w</sup>ong* could really serve as the phonetic in 通 *t'ung* merely on account of the relationship of the vowel-sounds and the similarity of the final sound (-ng), without regard to the initial, it is clear that this *i<sup>w</sup>ong* could also have served very well as the phonetic to a *kung*, or *k'ung*, or *ki<sup>u</sup>ng*, etc. We should therefore expect, if 甬 was used in a whole series of composed characters, to come across words like *t'ung*, *kung*, *sung*, *lung*, *mung*, etc. — all kinds of initials; but the series we actually find is this:

甬 phonetic in

甬  $d^wung$ , 通, 桶, 痛, 捅  $t^wung$ , 誦  $z_i^wong$  etc.

We see that the derivatives all begin with *dentals*; there is no instance of labials or gutturals. If we take another series, we find 爰  $j_i^wvn$  phonetic in:

諼  $\chi_i^wvn$ , 醵  $\gamma^wan$ , 緩  $\gamma uán$  etc;

— nothing but gutturals: no dentals or labials.

This cannot possibly be mere chance. The only imaginable explanation is that the phonetic 甬  $i^wong$  once upon a time began with a dental, e. g. *d-*, which was afterwards dropped, before the 6th century, and we can very well understand why a  $d_i^wong$  was chosen as the phonetic in 通  $t^wung$ , 甬  $d^wung$ , etc. (but not in *kung*, *lung*, etc.). This is then in entire conformity with the rule: agreement between the phonetic and the derivative in both initial sound and final sound. In the same way 爰  $j_i^wvn$  began with a *g-*:  $gj_i^wvn$ , which explains why it was used exclusively in words beginning with a guttural: 諼  $\chi_i^wvn$ , 醵  $\gamma uán$ , etc.

The question remains: Though it is clear that it must have been a lost dental in the former case, and a guttural in the latter, why was it just *d-* and *g-*? Here there are many interesting reasons. When we reconstruct the language of the 6th century, in a previously described manner, we arrive at the surprising result that it had

$|k|k^c|g^c|ng| \quad |t|t^c|d^c|n|$

but no common unaspirated *d-* and *g-* sounds. By means of our study of the 'phonetics' we get the explanation of

this peculiarity. The pre-Christian language had the whole series, but the unaspirated voiced stops dropped out:

| *k* | *k'* | *g* | *g'* | *ng* |    | *t* | *t'* | *d* | *d'* | *n* |    "

The *g*- was lost in cases like 爰 *j<sup>w</sup>on*, and *d*- disappeared in cases like 笛 *i<sup>w</sup>ong*. This conclusion receives another, striking support. When a word without an initial consonant sound (for I do not reckon *i*- or *ji*- as consonant sounds in the sense that *p*-, *p'*-, *t*-, *d'*-, etc. are consonants) serves as the phonetic to another word that does begin with a consonant, the former always has an *i*- or a *ji*-. In other words: cases of the type 笛 *i<sup>w</sup>ong*: 通 *tung*, or 爰 *j<sup>w</sup>on*: 緩 *yuân* are common, but cases like *ân* phonetic in *kân*, or *uân* phonetic in *tuân* are practically non-existent.

We are therefore forced to conclude that the *i*- or *ji*- in the phonetic reveals a lost *d*- or *g*-. This is in line with common phonetic phenomena. We have many examples in Swedish of a *d*- or *g*- having been lost before *i*: *d<sup>i</sup>up* and *g<sup>i</sup>uta* are now pronounced *iup* and *iuta*. And in the N. of Germany *g*- itself has become *i*-, as in *gans* when pronounced *i<sup>ans</sup>*. It is therefore but natural to imagine the process of development to have been from *d<sup>i</sup>wong* to *i<sup>w</sup>ong*.

In a similar manner I have been able to show that if 白 *b<sup>w</sup>ok* is the phonetic in 怕 *p<sup>a</sup>*, it is because in pre-Christian times there was a *-g* in 怕: *p<sup>a</sup>g*, which was dropped later. In other cases a *d*- was lost: when, for instance, 列 *liät* is the phonetic in 例 *liäi*, it is because *liäi* has developed from a pre-Christian *liäd*. I will not

tire my readers by setting forth the proof here, but may mention that, thanks to the 'phonetics' and their part in the composed characters, I have succeeded in many important points — some have not been touched on here — in drawing conclusions regarding the consonant system a couple of centuries B. C., and that there are hundreds of common words about the then pronunciation of which we can get a fairly concrete idea in this way. And I need scarcely say what an advantage it is for the linguist, if in this manner we can get back to a stage when the pruning process had not proceeded so far, when the words were still phonetically richly developed, a *p'ag* instead of a later *p'a*, a *dī<sup>w</sup>ong* instead of a later *i<sup>w</sup>ong*, a *diag* instead of a later *iäu*, etc. As soon as we reach the task of a serious comparative Indo-Chinese philological investigation, these pregnant wordbodies will be incomparably easier to work with than their later, mutilated descendants.

When we have done all we can with the purely Chinese material, the next great step will be, of course, comparative Indo-Chinese philological work. There is still a group of native material, it is true, which I have not yet mentioned, and which many Chinese scholars had set great hopes on during the last few centuries, but as to which I am none too optimistic. This is the rimes in the oldest poetry. Among the Confucian classics, one of the most important is Shī-king, the Canon of the Odes, 300 songs, all of which, with the exception of five that are said to be older, go back from the twelfth to the seventh centuries B. C., and Chinese philologists have been busily occupied in dividing up the rime-words of these songs into classes, in order to find a key to the pronunciation of the endings in this oldest Chinese. The difficulties, however, are many. In the first

place we know that a great number of these songs come from various small states in the China of those times, and dialectal differences must have played a part; so it is impossible to treat these rimes as homogeneous material. Furthermore, we have no idea how exacting poets were in their rimes. The songs in question are to a great extent simple folk-songs, and our experience from, let us say, the Scandinavian and English ballads teaches us that even very approximate rimes suffice in this class of poetry. The Swedish folk-songs, for instance, have *in* and *kring* in rime, *viter* and *liker*; the English ones have *dyke* and *knight*, *kin* and *him*, *man* and *won*, etc<sup>1</sup>. But as contributive material, a support for results obtained by other means, these rimes will certainly be of some importance in due time.

The next great task will therefore be the comparative study of Chinese, Siamese, Tibetan, and other languages belonging to the Sinitic family. The material for such an investigation might seem to be ready. We have not only a knowledge, through the alphabetical writing, of the pronunciation of Tibetan in the 7th century A. D., and of Siamese from the 1280'ies; we have also the pronunciation of reconstructed Chinese in the 6th century, and have been able to draw important conclusions about its consonants many centuries earlier. And yet the materials are hardly available in a fully satisfactory state. If Chinese had not become a cultural and literary language until about 600 A. D., like Tibetan, which, broadly speaking, only became a literary language from the 7th century onwards, the matter would be fairly simple. But during a long development of civilisation and literature, the meanings of words have been differentiated and often altered past recognition, and

<sup>1</sup>) See F. J. Child, *The English and Scottish popular ballads*.

it is therefore very risky to work with the Chinese words of the 6th century for comparative purposes, taking them with their meanings at that period. I have set forth earlier how, a few centuries A. D., Chinese obtained a very large vocabulary, at least some fifteen to twenty thousand words, simple monosyllables with their own characters, and therefore apparently primary words. I have pointed out that these must be but apparently primary words, and that the philologist must be able to refer these masses of words to a considerably smaller number of word-families, stems, and show how these masses of words branched off from the different stems, and what were their primary meanings. But until this is done, they constitute dangerous material. Owing to their numerousness and variations in meaning, they allow a comparative philologist to find parallels with almost anything. If he looks for something corresponding to the Tibetan *bud* 'dust cloud', he can find, for instance, in Kuang yün, a lexicon from the 11th century, based on our Ts'ie-yün from the 6th century, 23 different words that are all pronounced *b'uət* (the nearest sound-equivalent to *bud* that was possible with the sound-system of that time), 5 different *p'uət*, 20 different *piuət*, 9 different *b'iuət*, etc., and it is almost any odds that among all these there will be some word meaning 'dust' or 'earth' or 'dirt' or 'smoke' or 'powder' or 'cloud' or the like, which could without a blush be connected with the Tibetan *bud*. In this case the irony of fate decrees that we have not less than two words, 埴 *b'uət* and 拂 *b'iuət*, both meaning precisely 'dust cloud'. Embarras de richesse! Both are very unusual words, the fates of which are very obscure. They may perhaps represent a stem that is identical with the Tibetan *bud*, — perhaps not; for a similarity



in sound, discovered, not among two or three Chinese words, but among sixty odd, carries no proof with it. And yet, in the few attempts at comparative studies hitherto made, conclusions have sometimes been based on flimsier grounds than this.

In arranging words into families and stems, it is of decisive importance to get hold of the original meanings, and show the transition to their secondary, transferred meanings; here we again get help, and very valuable help, in the Chinese script, which in other respects puts such difficulties in our path. The many meanings of a word, though all deriving from a common fundamental sense, are often so divergent that it is difficult to see the development line; but here the characters often come to our aid, revealing which is the primary sense. Thus, we have a word that in the 6th century was pronounced *xjei* and means 'to hope'. We have one that was pronounced *xjei* and means 'valuable'. We have one that was pronounced *xjei* and means 'thin, rare'. If Chinese had had the ordinary alphabetic writing, it is not certain that we should have hit upon the idea of connecting these three words; but the Chinese script has the same character for all three, 希, and if we analyse this we shall find that it consists of 巾 'cloth, stuff' and 爻 a drawing of big meshes. The meaning that the originator of the sign wanted to indicate was therefore 'thin, loose, like loose stuff', and from this sense we then get the series: (1) thin — rare; (2) rare — valuable; (3) (as a verb:) to find scarce, valuable — to value — long for — hope for. For all these shades of meaning we can find examples in literature. Innumerable are the cases where in like manner the script gives a

valuable hint. 奚 *yiei* 'servant, slave' ought to be connected with the stem 系 *yiei* 'bind' (the sense being a bond, an unfree person). We have furthermore a word 燕 *'ien* 'swallow', originally a drawing of the bird, and and this enters as an element in 嚥 *'ien* 'to swallow'. That 'the swallow' is here not only the phonetic, but that the two words are etymologically identical, so that 'the swallow' got a name meaning 'the swallower, devourer', seems evident to me; we need only look at a brood of open-beaked young swallows being fed by their mother to get the impression that every chick is one great gape, or throat<sup>1</sup>.

If it is a chief condition for comparative investigations that the word-material be sufficiently worked up on the Chinese side, with regard to the actually existing stems and their fundamental senses, it is also at least equally important that conclusions be not drawn from conformities in sound in isolated cases. For then one could prove

<sup>1</sup>) By the way, why should it not be the same with the Teutonic word? The word 'the swallow' (Germ. *Schwalbe*, Swed. *svala*) has been derived in various ways, e. g. by Saussure, who connected it with Greek *alkuōn*, but no suggestion has met with general approval. In his new etymological dictionary of Swedish, Hellquist desists from giving any interpretation. My friends working in the field of Scandinavian philology now tell me that I might be quite justified in attaching to the primitive Teutonic stem *suelz-* 'to swallow' (Dan. *svelge*) a noun prim. Teut. *\*sualyōn* deriving from a still older *\*sualzyōn*. This would then have given O. H. G. *swalwa*. If the High Swedish *svala* has lost the original *-y-*, this remains, however, in Northern Swedish dialects in the form *svalv*, and *-y-* belongs to a suffix *-yōn* that is common just in bird-names, e. g. *sparrow*, Sw. *sparv*. (For these forms I have to thank my colleague Prof. Evald Lidén). In that case it is not by mere chance that the 'swallow' and to 'swallow' should be identical in sound in English — although they do not derive from identical older forms.

anything. One could cheerfully succeed in comparing quite a number of Indo-European roots with Chinese words, e. g. Ancient Chinese *t'iau* 'daylight, day' with I. E. *\*diēu-*, which we have in Latin *dies*, with exactly the same meaning. Or one could make Chinese a Teutonic language on the strength of cases like 行 *ɣvng* < *g'ǎng* = Icelandic and O. Sw. *ganga* 'to go', etc. In other words: conclusions must be drawn from *established sound-equivalences in series of words*. It may seem superfluous to the experienced linguist to underline this self-evident claim, but experience shows how easily this simple principle can be set aside unless a strict eye is kept on laxness in methods, in whatever branch of linguistics it may rear its head. It is perhaps natural, though not pardonable, to relax this claim somewhat in searching for kindred tongues for languages like Etruscan or Basque, but it is quite inexcusable in dealing with great and richly documented linguistic families like the Sinitic. Here we must positively, with strict methods, arrive at results that are absolutely sure.

## CHAPTER VI.

**I**n the foregoing we have seen how, in dealing with the older stages of the Chinese language and literature, we must handle material of very diverse nature. The main bulk has been purely native, to this have been added, as valuable contributions, foreign versions of Chinese words and Chinese forms of foreign words. The most important of the former were the three great categories of Chinese loan-words, Sino-Korean, Sino-Japanese, and Sino-Annamese; of the latter, the Indian and central-Asiatic words which in connection with the victory of Buddhism over East Asia, and the political expansion of China westwards, were incorporated in large numbers in Chinese literature in transcription with Chinese characters. Both these groups of material, however, are comparatively late, with their centre of gravity falling between the 5th and 9th centuries A. D., and they therefore do little more than support and check the reconstruction results we have arrived at with the help of the copious native sources, rime-lexicons, rime-tables, fan-ts'ie spellings and modern dialects.

But the investigator will not willingly be satisfied with this. He cannot forget the wonderful results arrived at by the study of loan-words in another field of study, the

Scandinavian languages, in that the oldest Scandinavian loan-words in Finnish have preserved for us valuable testimony of how the proto-Scandinavian tongue was pronounced, and thus complete the picture of that language as we have obtained it from other sources. And he argues that even if for older epochs we cannot find such quantities of words of foreign origin in Chinese, or of Chinese words in foreign languages, there must still, in all probability, be a small number of words here and there which may throw valuable light. For China was not hemmed in by deserts, but has always been surrounded by lively neighbours, with whom she has now been at war, now maintained friendly relations. It would be most unnatural if this had not left behind it some linguistic traces, from epochs that are considerably older than the great loanword periods. Let us look into the matter.

Our thoughts turn instinctively to the interesting realm to the east of China, Japan. But when we touch on this chapter, we open up the whole complicated problem of the origin of the Japanese, and their earliest history; since, before seeking to find very old Chinese words in Japanese, we must have some clear conception of what sort of language Japanese is, and how Chinese words can have entered into it.

The origins and early history of the Japanese race are wrapped in obscurity. That it is a mixed race seems to be certain, and it is generally allowed that it migrated into the country from abroad. The oldest Japanese myths, which are undoubtedly an echo of traditions from pre-Christian times, though they are only preserved in two texts from as late as the 8th century, *Kojiki*, published in 712, and *Nihongi*, published in 720 A. D., are localised to three different centres: one on the N. coast of the chief

island, one on its S. coast, and one on the most south-westerly of the Japanese islands, Kyushu. Of these, the ones from Idzumo in the north and Kyushu in the S. W. are the oldest, and according to Japanese tradition the Emperor Jimmu, the ancestor of the present reigning family, made a military expedition in 660 B. C. from Kyushu to the third centre, Yamato, and founded the dominion that afterwards became the first great centre in the development of the historical Japanese empire.

The traditions therefore indicate a first historical landmark on the north side, and one on the S. W. island, and it is suggested that while N. Asiatic tribes wandered in via Korea, seafaring colonists reached Kyushu. These are thought to have fused into the present Japanese race and gradually supplanted and of course partly absorbed the Ainu people, who of a certainty occupied a large part of the main island in olden times, though they now only exist as the scattered remnants of a race in the extreme north of Japan and in Saghalien.

For a confirmation of these theories we are referred to archæology and anthropology, and to linguistics. What the last-named has to inform us about, is whether Japanese is related to any other languages on the Asiatic mainland, or with any languages of the South Pacific, or whether it is a mixed language of these elements. The question is by no means easy to answer.

We must first bear in mind that the Japanese, wherever they may have come from, have been separated from their original kinsmen for at least two thousand years. In such a space of time, every language changes considerably. With regard to Japanese, we are very fortunately placed, for thanks to the Japanese phonetic writing, the Kana

script, we know exactly how the words were pronounced at least a thousand years ago. But for most of the N. Asiatic and the Austronesian languages we have only quite modern forms of words to go by, and with these it is somewhat hopeless to compare the Japanese words. Furthermore, as regards Korean, which many authors, with very slight proof, have declared Japanese to be closely related to, we do not even know how long it has been spoken in Korea. At the time of the birth of Christ there were a number of small states in Korea, and we know very little about their tribes and languages. Whether the present Korean was spoken in the peninsula in an early form in remote ages, or whether it was introduced by some migrating tribe from other parts of Asia in comparatively late times, we do not know. The investigation is not altogether hopeless, but it has a difficult and long-drawn task to face: the East Asiatic languages we know must first be studied comparatively, and parent languages must be reconstructed for such of them as are related to each other. The same process must be carried out with the Austronesian languages, or at least the parent languages must be reconstructed for some of the main groups in this mighty world of languages, the Indonesian, the Melanesian, the Polynesian groups. Then, and not until then, can we compare Old Japanese, such as it is preserved in, for instance, Kojiki, with these continental and Austronesian languages; and then, perhaps, we can get a complete answer to the question about the origin of the Japanese race.

To a certain extent, however, philology can already take up the problem now. In regard to one of the ethnical components, we are tempted to ask: Were not the south-western immigrants Chinese? Kyushu, the first halting-place

in their progress from the south, is the nearest of the islands to China, and it is indisputable that the Japanese and the Chinese both belong to the Mongolian race. As a matter of fact an attempt was made in the 1880'ies by an English Sinologist, E. H. Parker, to prove the affinity between Chinese and Japanese. He adduced a few hundred Chinese words which he considered could be identified with Japanese words. However, his method was so faulty, the similarities in sound were so thin and fortuitous, and the whole attempt so irresponsible that the distinguished Japanologist B. H. Chamberlain gave him a good dressing down. The idea of comparing Chinese and Japanese has unfortunately been miscredited through Parker, and has since then been practically shelved<sup>1</sup>.

The time was really not yet ripe for such a comparison, since one had not then the knowledge of Ancient Chinese that we have gained by reconstruction, as sketched above, of the Chinese of the 6th century. An investigation of this kind has now greater chances of succeeding.

Here we may at once state the fact that Chinese and Japanese are probably not related. The whole grammatical system, the formation of words, the phonetic type of words, etc., are as different as in Japanese and English. If the Kyushu-ites were dominant in the linguistic development, and one of the main elements in the Japanese race, they were not Chinese. But this in no wise prevents their language from having contained Chinese elements in the shape of loan-words supplied by cultural influence from China. Before the Japanese had begun, from the 5th century onwards, first via Korea and then directly, to import

<sup>1</sup>) J. Matsumura's big volume (1916) in the same style is even worse than Parker's attempt.



Chinese culture on a big scale in all departments of political and social life, they already had, as Kojiki and Nihongi prove, a not inconsiderable measure of culture. Had this developed entirely from within, or did it contain some elements derived from intercourse with the Chinese? That is the question. If such intercourse existed, it should be revealed linguistically in some detail or other.

Most authors who have written about primeval Japan, have considered, for completeness' sake, that they ought to mention the possibility of intercourse having existed with China in the earliest times, but there they generally stop. Thus Chamberlain, the same one who castigated Parker, has himself, in an article on the oldest stock of Japanese words, made some allusions to this, and has mentioned altogether 6 words which seem to him to point in that direction. An unkind fate will have it that 3 of them prove to be phonetically impossible. The question therefore deserves taking up afresh.

The historically documented route for the first cultural influence from China to Japan goes via Korea. From there was introduced the art of writing, according to both Kojiki and Nihongi, through a learned Korean, Wani or Wang-yin, who brought with him some Chinese classics, instructed the Japanese heir-apparent in them, and became the progenitor of a race of scribes. For this Nihongi gives a date which probably corresponds to 405 A. D. The Japanese may, of course, have made some acquaintance with writing in a more popular way at an earlier date, and this was only its official reception. On the whole, the above-mentioned works clearly show that already long before the great Chinesification period (7th century), masses of elements belonging especially to material culture had been

brought over direct by Koreans and Chinamen, chiefly skilled artisans, who had migrated to the islands and become the teachers of the Japanese. That the Chinese should have come exclusively via Korea, as is commonly believed, seems to me to be very improbable. Even after Yamato had become the political centre of the Japanese race, it kept a firm footing in Kyushu, and it is in Kyushu that was found the oldest memorial of the intercourse with China, viz. a gold seal from 50 A. D. From Kyushu to the opposite land, in the part that is now Shanghai, the distance is about the same as from the south of Greece to Sicily, a very feasible sail. From the Philippines a forcible gulf-stream runs north, Kuroshio 'the black salt', which strikes Japan just at the south coast of Kyushu, and Chinese seafarers who had ventured too far out to sea, could therefore easily be carried by the current to Kyushu. So everything seems to argue in favour of direct intercourse between E. China and Kyushu, intercourse that possibly goes further back than the one that passed over Korea, which latter can scarcely date back to pre-Christian times.

During the great period of Chinese influence, thousands of Chinese loan-words were introduced into Japanese, so that a very large percentage of the Japanese vocabulary has consisted ever since of Sino-Japanese words. These words were incorporated *en masse*, as literary loan-words according to fixed principles, and therefore they can easily be recognised even to-day, in spite of changes in sound they have undergone on Japanese soil, and be distinguished from the genuine Japanese word-material.

If, however, Chinese influence helped, much earlier still, to lay the foundation of the oldest Japanese culture, in the ways I have just suggested, we should be able to discover

among these other words, which have hitherto been regarded as genuinely Japanese, some that were borrowed direct and in a popular way, sharply differentiated from the later Sino-Japanese words, and possibly reflecting an older stage of the Chinese language.

In looking for such examples we must not lose sight of the fact that every language that borrows foreign words, refashions them to suit the normal pronunciation of its own. I need hardly refer to the manner in which French words have been mangled on being taken up into our Scandinavian languages. There is therefore as a rule a certain measure of sound-substitution, often a very radical one. In Japanese we have a possibility of guessing, a priori, on what lines these substitutions will proceed, for we need only examine, in this connection, the later borrowed Sino-Japanese words, on the one hand, and on the other the sound-resources in Old Japanese, as stored in Kojiki. Thus, we find:

(1) The Japanese could not reproduce *-p*, *-t*, *-k* at the end of words, but had to append a supporting vowel. Chinese *puk* became Sino-Jap. *puku*, *kat* became *katu*, *kiet* became *ketu* or *keti*, *kap* became *kapu*. This finds a striking parallel in Finnish, which has the same antipathy, so that modern Swedish loan-words receive a similar addition in the shape of a parasite vowel; Sw. *kräm* becomes Fi. *rämi*, etc. An examination of Kojiki shows that in the oldest Japanese no word at all could end in a consonant, so we must expect a parasite vowel after every Chinese final consonant. In the more scholarly Sino-Japanese they only used *-i* (after a palatal vowel) and *-u* (after a velar one); but we cannot expect such a scientific procedure in popular loans, and must look for a fluctuation between the existing vowels, thus:

$$kak \begin{cases} a \\ i \\ e \\ u \\ o \end{cases}$$

(2) Voiced plosives in the initial sounds, *b-*, *d-*, *g-*, simply did not exist in genuine Japanese words. It is true that we now find that in the Sino-Japanese Go-on version, in spite of this, Old Chinese *b<sup>h</sup>-*, *d<sup>h</sup>-*, *g<sup>h</sup>-* are represented by *b-*, *d-*, *g-*; but this refers, as I have said, to literary loans made by comparatively learned borrowers. But if at an early stage a Chinese *b<sup>h</sup>a* was borrowed in a purely popular, oral way, it goes without saying that *pa* was quite automatically substituted for *b<sup>h</sup>a* by a people that in its own language could say *pa* but not *ba*. We should therefore expect that Chinese *b<sup>h</sup>a* would become *pa*, *d<sup>h</sup>a* > *ta* etc.

(3) Several consonants together were impossible in Old Japanese, and the consonant group had to be simplified. A Chinese *klam* had to become *kam* or *ram* (*l* does not exist in Japanese); cf. Swed. *kräm* becoming Finnish *rämi*.

(4) In Chinese *i* and *u* occur abundantly in combination with other vowels, but in Sino-Japanese they are often dropped: A. Ch. *kiät* is S.-J. *ketu*, *kiu* is *ku*, *tuän* is *tan*. Or, again, only the *i* or *u* are kept, and the other vowels dropped, e. g. A. Ch. *kiĕn* > S.-J. *kin*, *kiuən* > *kun* — a very natural simplification in borrowing.

(5) Certain vowels occurred in Chinese which were lacking in Japanese. These were reproduced by other, proximal vowels. Sino-Japanese renders *ɐ* (Engl. *but*) by *a*, *ɔ* by *a* or *o*, *ə* by *o*: A. Ch. *kɐk* > S.-J. *kaku*, *kɔk* > *kaku* or *koku*, *kən* > *kon*.

Remembering these rules, deduced, as they are, from well-known facts in the Japanese language, we shall now examine a few words that we might take to be loan-words from remote antiquity. Firstly, two words that Chamberlain, for material as well as phonetic reasons, had already judged to be Chinese loans.

Japanese		Pekinese
<i>uma</i> 'horse'	=	馬 <i>ma</i>
<i>ume</i> 'plum tree'	=	梅 <i>mei</i>

As far as the latter word is concerned, this identification is a hit due to luck. Pekinese *mei* derives from a *mudi* of the 6th century, and that *mudi* cannot have become *ume* is clear. However, we have grounds for knowing that, a few centuries after Christ, in the part of China that faces Kyushu, the word was not pronounced *mudi* (as in N. China) but *mue* (with the same transition as Caesar [*kaisar*] to *cēsar*), which explains the Japanese *e* in *ume*. As to the *u-* in *uma*, *ume*, no explanation has hitherto been given of it, but I have been able to show, for reasons that I cannot enter into here, that the Ancient Chinese *m-* was pronounced with protruded lips, *<sup>h</sup>ma*, *<sup>h</sup>mue*, which is reflected in the Japanese loan-forms.

New Jap. Anc. Jap.

Pek. Anc. Chin.

<i>ie</i>	<i>ipe</i> 'house' =	邑 <i>i</i>	<i>·iəp</i> 'town, village'.
<i>sato</i>	<i>sato</i> 'village' =	室 <i>shī</i>	<i>śiēt</i> < <i>śiāt</i> 'house'.

In these examples we find the parasite vowels *e* and *o*, and in the former case *i* reproduced and *ə* dropped, in the latter case *i* dropped and *a* reproduced. In both the sense has been shifted. *·iəp* was originally a community of 24 families, and has since been used in the senses of 'large

town' and 'village'. The word *śiēt*, on the other hand, means, it is true, a 'house', but in the patriarchal sense of the common habitation for a great family, various branches of the same lineage, and therefore also means 'family, clan'. The Japanese *sato* (village) was probably not larger than a clan settlement. We must, moreover, remember that loan-words often get their meanings considerably shifted. A good example of this is the French *bureau*, which means 'writing-desk', but which both in Swedish and English has come to mean 'chest of drawers', which is in French *commode*.

If in primitive Japan the house had a Chinese name, we need not be surprised that 'to build, erect a wall' is

N. J.      A. J.                      Pek.      A. Ch.

*tsuku*    *tuku* 'build' = 築 *chu*    *t'iuk* 'build'.

and that the house was surrounded by a hedge, a planted defensive wall:

N. J.    A. J.                      Pek.    A. Ch.

*kaki*    *kaki* 'hedge' = 榻 *ko*    *kpk* 'partition, screen'.

If the Chinese were instructors in carpentry, that is probably the reason for 'to split, cleave' being called

N. J.    A. J.                      Pek.    A. Ch.

*saku*    *saku* 'cleave' = 析 *si*    *siäk* < *siak* 'cleave'.

To the house belongs furthermore

N. J.    A. J.                      Pek.    A. Ch.

*yuka*    *yuka* 'floor' = 闕 *ü*    *i'ək* 'threshold'.

Now it might be argued that 'floor' and 'threshold' are by no means the same thing. But we must remember that in the earliest times the Japanese house probably had no complete floor, and that *yuka* only indicated a wooden foot board along the walls. It would then actually mean

'sill', and it is significant that in English the word 'sill' means not only 'horizontal timber serving as the foundation of a wall', but also 'threshold' (cf. German *Schwelle*) and sometimes even 'floor' (see the Oxford English Dictionary). The correspondence in sound, *yuka: i<sup>w</sup>ək*, is certainly quite remarkable.

Two objects belonging to the material culture that might a priori be expected to be loans are silk and rice.

N. J.	A. J.		Pek.	A. Ch.
<i>kinu</i>	<i>kinu</i>	'silk stuff' = 絹	<i>küan</i>	<i>k<sup>w</sup>än</i> 'silk stuff',
( <i>ine</i> )	<i>sine</i>	'rice' = 秬	<i>sien</i>	<i>siän</i> 'rice' (of a certain kind).

The parasite vowels are *-u* and *-e*, and *i* has supplanted the other vowel elements. That *kinu* and *k<sup>w</sup>än* are identical, was already suggested by Parker. Rice is normally *ine*, but Kojiki has an alternative form *sine*, of which *ine* may well be regarded as the truncated variant. It is also possible that we have a Chinese stem in the word for silkworm, N. J. *kaiko*, A. J. *kapi-ko*. *Ko* only means 'child, the young of an animal', and thus we get

N. J.	A. J.		Pek.	A. Ch.
<i>kai</i>	<i>kapi-</i>	= 蛱	<i>kie, kia kap, kiek</i>	'butterfly'

and *kapiko* would mean then simply 'butterfly-child', i. e. caterpillar. The difficulty is here that 蛱 in Chinese means 'butterfly' generally, while the silkworm since time immemorial has had a specific name, a technical term (*tsan*), so that the identification must be taken cautiously (for 蛱 the Kuang-yün gives the reading *kiep*, but many dialects, e. g. Cantonese, indicate clearly an ancient *\*kap*).

To the vegetable kingdom belong furthermore two typic-

al cultural loan-words, and a third botanical term deserves at least to be discussed in this connection.

N. J.	A. Jap.		Pek.	A. Ch.
<i>take</i>	<i>take</i> 'bamboo'	= 竹	<i>chu</i>	<i>t'iuk</i> 'bamboo',
<i>mugi</i>	<i>mugi</i> 'wheat'	= 麥	<i>mai</i>	<i>m<sup>w</sup>ɸk</i> 'wheat',
<i>sugi</i>	<i>sugi</i> 'Cryptomeria'	= 松	<i>sung</i>	<i>ziung</i> 'Pinus'.

We see that *take* and *t'iuk* correspond poorly in vocalism. But I have the greatest difficulty in believing that the bamboo, one of the chief cultivated plants of E. Asia, used for practically every purpose, did not get its multifarious services taught by the Chinese instructors to the primitive Japanese, and that therefore *take* and *t'iuk* are the same word. We must remember that *t'iuk* was the pronunciation in the 6th century A. D.; possibly the form *take* reveals to us that many centuries earlier Chinese had a more open vowel in this word.

The words *mugi* and *sugi* have this in common—that they both end in *gi*, a syllable that ends a number of Japanese words for varieties of trees, bushes and plants. *-gi* is the word *ki* 'tree' with the sonorizing that is common in Japanese (so-called *nigori*), of the intervocal consonant to *-gi*, e. g. *yane* 'roof', *yanagi* 'roof-wood', i. e. 'sallow', etc. Chamberlain has rightly pointed out that in both *mugi* and *sugi* we have this *-gi*. So we must regard them as contracted forms of compounds: *muku-gi* to *mugi*, *sugu-gi* to *sugi*. As to the latter, we must substitute *s-* for the *z-* which does not exist in genuine Japanese, and similarly substitute *-g(u)* for the *-ng* that does not exist. Both cases thus fit into the substitution laws. It should however be remarked that *sugi*, as Chamberlain has pointed out, may



possibly go back to a stem *sugu* 'straight', i. e. *sugu-gi* 'the straight tree'.

The intervocal sonorizing of *p*, *t*, *k*, to *b*, *d*, *g*, which is so common in Japanese, perhaps explains the forms of two words belonging to material culture which may be suspected of being loan-words:

N. J. A. J.

Pek. A. Ch.

*togu togu* 'grind, polish' = 琢 *cho t'āk* 'grind, polish',

*hagu pagu* 'flay' 剥 *po pāk* 'flay'.

The common feature in these, both of which have sonorizing, is that they are verbal stems. The cutting of precious stones is an art that may be suspected of deriving from the mainland, and so, too, is the leather industry and technique.

Another industry that is exceedingly old in China, and the technique of which was probably taught the old Japanese, is obtaining salt from sea-water. Right into our own times the usual method in Japan has been the one common on the coasts of China, viz., by enclosing the water in shallow dams along the coast, thus forming a *salt marsh* or *swamp*. That is why we may try to identify

N. J. A. J.

Pek. A. Ch.

*shio sipo* 'salt' = 涇 *shī śiəp* 'moist, marsh, swamp'.

Two words of administrative character, which Parker was quite correct in identifying, are:

N. J. A. J.

Pek. A. J.

*kuni kuni* 'principality' = 郡 *kūn g'iuən* 'prefecture',

and we need only point out that the case agrees with the substitution laws we have drawn up (*-i* is parasitic, Ch. *g'*- is reproduced by *k*-, and only one of the vowels is represented).

Among household furniture we find a couple of very possible examples.

N. J. A. J.

Pek. A. Ch.

*fune pune* 'vessel':= 盆 *p'én b'uən* 'vessel',

*kama kama* 'pot' = 罐 *k'an k'ām* 'pot'.

In later Japanese, *pune* means exclusively 'boat', with a transition well illustrated by the E. 'vessel' = 'ship'; but in Kojiki it means 'vessel', so the case is a very likely one. *Kama* suffers from the difficulty that there is no real common word *k'ām* 'pot' in Ancient Chinese literature. But that a popular *k'ām* of very wide distribution must have existed, is evident from the fact that the lexicons give a whole row of characters pronounced *k'ām* and with this meaning: 罍, 甕, etc. Evidently all sorts of scribes wanted to write *k'ām* 'pot', and for lack of a classical character each made up his own.

Returning to agriculture, we find an interesting case in:

N. J. A. J.

Pek. A. Ch.

*kama kama* 'sickle' = 鎌 *lien liām* 'sickle'.

This identification may seem bold. But the fact is that the Chinese word 鎌 is written with a phonetic 兼 A. Ch. *kiem*, which shows that a *k-* or *g-* had been dropped, so that *liām* derives from an older *kliām* or *gliām* (or possibly *klam* or *glam*), which according to our substitution laws should give an A. Jap. *kam/a*, and the example seems to me to belong to our most interesting ones.

If the Chinese were instructors in agriculture, it is not to be wondered at if the name for 'summer', the season of vegetation, is of Chinese origin:

N. J.      A. J.

Pek.      A. Ch.

*natsu*    *natu* 'summer' = 熱 *jē*      *n̄ziāt* < *n̄iat* 'hot, heat', an identification already proposed by Parker.

All the words I have cited have a Japanese form that differs decidedly from the sound-type of the later Sino-Japanese words. From what time they date, it is of course impossible to say. Some of them, e. g. *uma* 'horse' and *kuni* 'land' may be quite late, even as late as the beginning of the Chinesification period, thus popular loans, parallel and contemporary with the earliest Sino-Japanese words. Others, on the contrary, must be much older. Cases such as *sato* 'village' = *śiēt* < *śiūt*, *natsu* 'summer' = *n̄ziāt* < *n̄iūt*, and above all *kama* 'sickle' = *liām* < *kl(i)am* or *gl(i)am* point forcibly to a much earlier stage in the Chinese sound-development, the last case perhaps even to pre-Christian times.

I have already touched in detail on the difficulties that are presented when words from other languages are compared with Chinese words. I pointed out that, thanks to the mass of apparently primary words, a diligent employer of the lexicon can take almost any simple foreign word, and find in Chinese something very similar in sound and meaning. But he must then often be satisfied with the most unusual words the Chinese lexicon has to offer. The comparisons I have just made, do not suffer from this weakness, for with the exception of *k'ām* 'pot' — the documentary evidence of which is so slight in literature, but a word that was evidently in popular use in olden times — all the words in question were fully current in Ancient Chinese, and most of them actually belong to the élite of the vocabulary. The Japanese words, again, are all to be found in *Kojiki*, and belong to the very commonest words

of the later Japanese language. All the examples touch concrete things, associated with material culture, and many are such as give us good grounds for suspecting Chinese importation. With one exception (*take: t'uk*) the correspondences in sound are satisfactory, viewed in the light of the substitution laws set forth. It may therefore appear to be all on a very safe basis, and that we can almost speak of proof. Yet I will not do this at the present stage. An important link in the chain of proof is missing, a link which at present cannot be supplied, and the absence of which every linguist will at once notice: absolute proof cannot be gained until we have shown that no other languages, possibly cognate with Japanese, possess forms that can, as well as or better than Chinese, explain the Ancient Japanese forms. I will show by an example what I mean.

The ordinary Chinese word for 'bear' (ursus) is Mandarin *hiung*, developed from a 6th century *ɣiung*, which in its turn, as I have been able to prove, developed from a still older *g'iung*. This may, however, correspond just as well to an archaic Chinese *g'ium* as to *g'iung*; for if, in pre-Christian times, there was a form *g'ium*, this would, according to known laws in Chinese, inevitably become *g'iung* by dissimilation: a labial final *-m* was impossible after labial main vowel *-u-*. A parallel shows this. The word for 'wind' was in pre-Christian times *pium*<sup>1</sup>, and became Anc. Ch. *piung* by such dissimilation. That 'bear' really was a *g'ium* and not a *g'iung* in pre-Christian times is shown by a form in the Swatow dialect, the most archaic, most peculiar of all known Chinese dialects. There, 'bear' has

<sup>1</sup>) That this was the case is shown by the fact that the character for 'wind' 風 has as phonetic 凡 A. Ch. *b'ium*. Also, in the old Shī-king, the word constantly rimes with words ending in *-m*.

a form *him*, which is very precious, for it has kept the archaic *-m*. In Japanese 'bear' is *kuma*, and it is very tempting to swell our list with the following equation:

N. J.      A. J.                      Pek.      A. Ch.

*kuma*      *kuma* = 熊      *hiung*      *γiung* < *g'iung* < *g'ium*

The case is in accord with the substitution-laws, the meaning is quite clear and concrete, and as surely as bearskin was ever a desirable commodity and object of trade, as easily can we imagine a Chinese loan-word here. Now, however, if we go to Korea, where the Japanese made dreadful ravages as vikings during the first few centuries of the Christian era, and the language of which country is held by many to be related to Japanese, 'bear' is there called *kom*, and it is more natural to connect Kor. *kom* and Jap. *kuma*, either as words deriving from a common ancestor, or as a Korean loan in Japanese, than to propose the Chinese *g'ium* as a loan-word in Japan. How, in its turn, it can be that 'bear' is *g'ium* in Chinese and *kom* in Korean, if this is more than a mere chance, is a problem to be solved in the future, and has nothing to do with the question of the Japanese loans. However, we see what dangers beset the philologist, and until all the languages suspected of relationship or alliance with Japanese have been included in the investigation, which will be an extensive piece of work for especially the Japanologists — the examples I have given are of course only a gleanings, and many other could have been adduced — I can only regard my speculations about Chinese loan-words in Japanese as a theory, even if at least some of the identifications appear to me to have a good deal in their favour.

If from the East we turn our looks to the North and North-west in order to see if an occasional word from, let

us say, the Turkish languages from a period older than the great epoch of Central Asiatic connections during the first thousand years A. D. can be traced, we are on safer ground. In the dynastic annals of the history of China from the first and second centuries B. C., we find several Hunnish words transcribed with Chinese characters — the Huns were the ancestors of the later Turks. To take an example. Some twenty years ago, the excellent German-American Sinologist Fr. Hirth pointed out an interesting passage from 47 B. C. The Hun prince was making a treaty with the Chinese, and, as it is put, stirred the sacrificial cup with a *king-lu*, which is asserted to mean 'knife'. Hirth stated that *king-lu* in older times was pronounced *king-luk*, and identified the word with *kingrak*, which occurs in various Turkish dialects, and means just 'knife'. It has been objected that this 路 *lu* never had a *-k* at the end, this being an invention of Hirth's. The S. Chinese dialects, as we know, preserve old *-k* at the end of a word, but in Canton this 路 is *lou*, without *-k*, and in the language of the 6th century as we have reconstructed it, *king-lu* was *kieng-luo*. Nevertheless, Hirth was on the right track. For 路 is precisely one of those words in which, because of the phonetic it contains, I can infer a *-g* in pre-Christian times, (cf. p. 112 above) and a *kieng-luog* corresponding to *kingrak* is a considerable improvement.

I will not give a list of Old Turkish words in Chinese literature, but will confine myself to one more example. In this I shall attempt to rehabilitate a word that our etymologists have treated in stepmotherly fashion, a word that is perhaps the most international in the world, next to the Chinese word *tea*.

If we look up the word *arrack* in the current lexicons, we shall find that it is explained, e. g. in Yule, in his great lexicon Hobson-Jobson, as derived from an Arabic word '*arak*', which means 'perspiration', from which it would have to mean 'exudation', 'sap drawn from the date palm', then broadly 'spirits'. From there it is supposed to have spread in different variants of form to Syria and Egypt, to various parts of India and the East Indian islands (in Malay *arak*), and then to crop up as a modern loan-word in every European language, and hence also in America. These statements, however, do no justice to the enormous distribution of the word either in time or space. The fact is that it can be traced back to a far earlier stage than Arabic literature touches, viz., the Hunnish language, via Chinese literature. The Chinese general Li Ling, who deserted to the Huns, wrote in 91 B. C. to his friend Su Wu and mentioned a drink 酪 A. Ch. *lák* (deriving from an archaic *glák*) — Chinese had no *r*-, so that a *rak* had to be rendered by *lák*. This *lak*-drink, a forerunner of the later kumyss, an alcoholic drink of fermented milk, also came into vogue in China. It was given up there long ago, and hence the word *lák* (Pek. *lo*) ceased to exist as a living word outside literature, but it occurs in the whole of N. Asia in the sense of 'spirits'. Radloff, in his big Turkish lexicon, has given us a whole series of forms: *arakı* in the Altai dialect, Teleutic, Sagaic, Kazan, Tobol, Crimea dialects; *rakı* in Osmanli; furthermore we have *arıkı* in Jakutic, *arki* in Manchu; the word also occurs in Tibetan under the form '*a-rag*'. But this is not all. Even Japanese has the word, and in a form which shows that it entered the language very early. For there it is *sake*, an interesting case of sound-substitution. In the Kojiki

language, words cannot normally begin with other consonants than *p-*, *t-*, *k-*, *m-*, *n-*, *s-*, while *r-* was introduced first with the Sino-Japanese loans. So to reproduce a foreign *rak*, *sak* plus supporting vowel was the nearest equivalent; hence *sake*.

The cope-stone of the structure is furnished by Ainu, a language that entered the Japanese islands in pre-historic times. Even there the word exists, in the form *arakke*, which shows that the Ainus did not get it from the Japanese, but from elsewhere. We thus see that in the stem *arak*, *rak*- we have a Central and N. Asiatic word which already in pre-Christian times reached all the races of the extreme Orient, was incorporated in Chinese, and conquered the whole world via Arabian.<sup>1</sup>

<sup>1</sup>) For further details see my review in *Deutsche Literaturzeitung* 1926 of Conrady's article *Alte westöstliche Kulturwörter*.



## CHAPTER VII.

As an imposing monument from hoary antiquity, overshadowing the whole East-Asiatic cultural world, and in no wise a ruin, but a monument still fully complete in its majesty, stands the Chinese script and Chinese literature, as the expression of the Chinese genius. Deeply loved and honoured by hundreds of millions of souls, this writing and literature has such a firm position that only mighty powers could shake it. And yet, coming generations may perhaps witness that this Old Chinese, still alive to-day, may be removed from its dominating position in the centre of the East-Asiatic communities, and put away as a museum curiosity, a memento of past ages. Perhaps; for destructively powerful are the forces at work in present-day China that call for change, for a levelling with the nations of the West, for simple and practical methods, even in the world of books, in order to ease the path towards democratism, the great panacea of the present age.

From time to time we read in our western newspapers paragraphs about the Chinese having now created an alphabetic spelling, so that they can discard their old, idiotic script, and at last become a civilised people like the rest

of us. And with a sniff we think, «well, it was about time; what an extraordinarily unpractical race not to have thought of that before!»

But when he reads such news, the western wiseacre little suspects as a rule that he has glimpsed but a slight symptom of a great crisis that East Asia — by no means China alone, but also to an almost equal extent, Japan — has impending, and which has already begun; a crisis in which gigantic cultural forces are at grips with each other in a struggle that will probably last for many decades, and the issue of which is uncertain. Let us briefly consider a few aspects of this involved problem, taking Japan first and then China.

When Japan borrowed Chinese culture in every sphere, especially from the 7th century onwards, it also took over, as I have mentioned, masses of Chinese words, and in addition the Chinese writing. It was a great cultural gift that China thereby gave to a country which had probably then no written language, but, owing to the incapacity of the Japanese to use the gift wisely and systematically, it became a veritable snare. As things turned out, the Japanese writing is now the most complicated and unpractical to be found in the world, incomparably worse than the Chinese, which is systematic, unambiguous, and in many respects excellent.

What made it go wrong from the very beginning was that the Chinese loan-words came from two different parts and in two soundversions. The same Chinese sign 平, A. Ch. *l'iang*, which means 'even, flat', was borrowed in the Kan-on version from N. China in the form *pegi*, which soon became *pei*, and finally *hei*; in the Go-on version from E. China *b'iang* became *biagu*, which soon became *biau*,

and which is now *biō*. For the character 平 we have thus in Japan two Sino-Japanese readings, which have nothing in common, *hei* and *biō*, one being used in *one* phrase, the other in another, and you have to learn each case by heart. But that is not all. Not only did the Japanese use the Chinese characters for the Chinese words they borrowed, but they also applied them to their own native, genuine Japanese words, for which they presumably had no written notation up to then; and so this same 平 'flat' was made to serve for the Japanese word *taira* with the same meaning 'flat, a plain'. But as there was furthermore a genuine Japanese word (cognate with the former) *hira* meaning 'flat, a plain', this, too, was written with the same 平. When in a Japanese book you come across the character 平, you cannot tell a priori whether to read it *hei* or *biō* or *taira* or *hira*; you must be an experienced reader and conversant with the phrase in which the word occurs, to know how it is to be read. In this way it is not unusual for one and the same character, in addition to several Sino-Japanese readings, to have a whole string of Japanese readings. And conversely: if we take the Japanese word, we find the word *hada* 'skin' reproduced by one author with the Chinese character 肌 (in Chinese pronounced *ki* 'flesh, bare flesh'), by another with the Chinese character 膚 (in Chinese pronounced *fu* 'skin'). Or it may happen that the same author in one place uses 肌 for *hada*, and then a few lines further on and on the same page uses 膚 for the same *hada*. According as an idea has many synonymous words in Chinese, and many Chinese char-

acters for them, we see that one and the same Japanese word for this idea can be written with a whole series of different Chinese characters, as the fancy takes the writer.

But this is not all. When it is not a case of merely single words being represented by Chinese characters of the same meaning, but groups of several words, the Japanese sometimes write, not the direct Chinese equivalent, word for word, but a free translation. We have, for instance, the word *ha-dome*, really 'tooth-stopper' i. e. 'brake on a wheel'. This is written by some authors correctly:

齒 止

tooth - stop

but by others

制 動 機

regulate - movement - apparatus

consequently a Chinese translation, which does not answer to the stems that make up the word, *ha* 'tooth' and *to-meru* 'to stop'. Sometimes we can see really comical — and yet generally accepted — extravagances of this kind. Thus, the word *urusai* 'troublesome, annoying', a genuine Japanese unit, reproduced by three Chinese characters 五月蠅 (*fifth-month-fly*) 'flies of the fifth month' (the middle of the summer, according to the Chinese calendar). This is indeed a very picturesque description for something 'troublesome', but a very free rendering of the Japanese adjective *urusai*.

The Japanese language is of quite a different type from the inflectionless Chinese. It has a richly developed system of inflections in its verbs, and the Chinese writing could not therefore be used with advantage for Japanese purposes. So in the 9th century they had already con-

structed a phonetic script in Japan, the so-called Kana, which by means of 50 different symbols can reproduce every possible sound in Japanese, e. g. か な *ka-na*, ひ と *hi-to* 'man', etc. These symbols are really shortenings of Chinese characters, which were used as pure phonetic symbols in Old Japanese, e. g. か from Chin. 加 A. Ch. *ka*, etc., and in order to make the matter a little more pleasant and complicated, they appear in two, graphically quite different systems, Katakana and Hiragana, something like the German and Latin character variations of our alphabet. They are used preferably for endings and auxiliaries; while the pregnant words, substantives, adjectives, verbs, adverbs, whether they are genuine Japanese or Sino-Japanese, are regularly written with the Chinese symbols in the hopelessly inconsistent way I have just described. But because it is so excessively difficult to manage to read the latter, you often see, especially in print intended for simple readers, and therefore as a rule in the newspapers, that the pronunciation of the Chinese symbols is indicated by small Kana at the right hand side of each character. Here is an example with Roman lettered substituted for the Kana signs:

*ka*

今 <sup>*i*</sup>  
*ma*

*ra*

*wa*

出 *de*

雨 <sup>*a*</sup>  
*me*

*ma*

*ga*

*se*

降 *fu*

*n*

*ru*

*Ima wa ame-ga fur-u kara de masen*  
Now as-to, rain fall-s because, (I) go-out do-not.

Here an attentive observer cannot help asking: If the Japanese already have a system of phonetic writing, which is fully effective, and not only shows them endings and auxiliaries, but also indicates the pronunciation of the chief words, why do they not get entirely rid of the Chinese writing, and write exclusively in Kana, or simply a transcription of Kana with ordinary letters?

It is clear that such a reform would meet with determined resistance before it conquered. Why, it would mean an almost unparalleled rupture with all the old tradition, the clerical customs of fifteen hundred years, and the old literature would at a blow become a foreign world for the children of the coming generation, with their phonetic script. When in our own countries we see what an outcry against and determined resistance to even quite unimportant spelling reforms arises, and how impossible it is, for instance, to get Englishmen to make the least modification in the crazy English spelling, one is inclined to think that a reform in the direction of purely phonetic writing in Japan is not to be hoped for. But we must remember that the conditions are in one respect quite different. The present situation in Japan is unendurable, which is not the case in any western country. And so, as necessity knows no law, we might expect, in the abstract, that a reform is psychologically possible. Alphabetical writing would, of course, be the most radical and offensive, and it would perhaps be difficult to carry it through; but it is unnecessary. The simple Hiragana system, with its easily written 50 symbols, is well suited to the phonetic structure of the Japanese language, works excellently with a few modifications, is easy for foreigners to learn, and would therefore, from an international point of view, do

just as well as Russian or Greek characters. To carry through the pure Kana style would not be amiss, if the language permitted it; but, as we shall see, this is by no means certain a priori.

The deciding factor is therefore not the psychological one, but quite another: could we understand Japanese written entirely on phonetic principles, entirely in Kana or in Latin letters? Here we cannot answer with a simple yes or no. We must distinguish between the two elements of the language, the genuine Japanese and the Sino-Japanese words.

As to the former, there is no reasonable argument against their being written phonetically. Japanese is a polysyllabic language with quite clear and easily distinguishable words. There are, certainly, a number of homophones, but scarcely more than in a western language, and Kana or letters would do excellently for this word-material. The same is true of a small number of the Sino-Japanese words, e. g. *niku* 'meat', which have become so thoroughly incorporated in daily speech that they are fully comparable with genuine Japanese words. If the Japanese spelling reformers could open their eyes to this fact, that the abolishing of the Chinese characters for all genuine Japanese words and for the perfectly Japanified everyday loan-words is possible and easy, the worst trouble in the Japanese writing would be eliminated at a stroke. One would have to go on, of course, learning many Chinese characters, but only one, or at the most two readings (Kan-on and Go-on) for each, whereas now we have to learn in addition a whole string of genuine Japanese readings attributed to this character; and similarly the present confusion of having two or more Chinese characters

for one genuine Japanese word, would disappear entirely. We should escape the present doubt about whether a character we meet with in a text is to be read as Sino-Japanese or as genuine Japanese. In other words: more than 50 % of all the troubles caused by the Japanese writing would be eliminated at a stroke by a reform that could be carried out *to-day*.

So far, well and good. But this does not separate all the chaff from the grain. The colloquial language, which shall be immediately clear to the listener's ears, has a preponderance of genuine Japanese words, or words that have been completely Japanified. A speech in such language could therefore be written down phonetically with Kana or letters, and could be read from the written page and understood. But as soon as the style becomes in the least exalted, or somewhat more abstract ideas are introduced, or things *literary* are dealt with, or even in the journalistic style, the word-material at once alters. There comes a heavy admixture of Sino-Japanese words, and this brings us into a quandary. These, too, *can*, it is true, be written phonetically with Kana or letters; but if in Ancient Chinese, from which these words were taken, the number of homonymous words was already so great that the colloquial language had to go on different lines from the language of literature, in order to be comprehensible, as we have pointed out; it follows that the poverty in sounds, and the quantity of homophones must be equally great in the borrowed Sino-Japanese. In fact, the trouble is intensified, for on Japanese soil, in part already in the process of borrowing and in part by the evolution during centuries, many syllables that were different in Ancient Chinese have become identical in modern Japanese pronunciation:



Anc. Chin.		Anc. Jap.	N. Jap.
<i>tsiāng</i> , <i>ts'iang</i> , <i>siāng</i>	have all become	<i>siagu</i>	<i>siau</i> > <i>shō</i>
<i>tsiāng</i> , <i>ts'iang</i> , <i>siāng</i>	»	<i>siagu</i>	<i>siau</i> > <i>shō</i>
<i>tsiāng</i> , <i>ts'iang</i> , <i>siāng</i>	»	<i>siogu</i>	<i>siou</i> > <i>shō</i>
<i>tsieu</i> , <i>ts'ieu</i> , <i>sieu</i>	»	<i>seu</i>	> <i>shō</i>
<i>tsiāp</i> , <i>ts'iap</i> , <i>siāp</i>	»	<i>sepu</i>	> <i>sewu</i> > <i>shō</i>
etc.			

As a matter of fact, things have reached such a pass that while Pekinese has at least 420 syllables among which its thousands of words are distributed, and while moreover the four tones do much to differentiate the words, Sino-Japanese has not more than about 250 forms corresponding to at least 3000 simple loan-words that are fully current in the simpler literary style. In a little lexicon that includes 3000 Sino-Japanese characters, I find 70 words that are all pronounced *shō*. It is true that the Kana spelling is conservative and distinguishes many words that have become identical in pronunciation, so that of these 70 *shō* words 29 are written *siau*, 14 *siou*, 24 *seu* and 3 *sepu*: but the impossibility of keeping these words apart by hearing alone (or with the help of only phonetic script) is nevertheless equally absolute and even more absolute than in China. I will allow that they occur mostly in compound words, specially coupled by twos, but not even this helps, for with the enormous number of syllables that have become identical, too many homophones have arisen even in these compounds. Thus, for instance, a much used Japanese lexicon (by Inoue) includes 12 different compounds *shō-kō* (written with different characters), 4 different *sho-kō*, 7 different *shō-ko*, etc. In other words: an average Japanese book or newspaper text, which always includes a very big percentage of Sino-Japanese

words, cannot be read aloud and understood without the help of the eye. It can consequently not be written with only phonetic symbols, but Chinese characters must be used as well.

The conclusion to be drawn from this is, again, of necessity: if the Japanese are to be freed from Old China, lying like a nightmare over Japan in the shape of its script, they must get rid of the Sino-Japanese stock of words in all cases except those relatively few instances where a word has been fully assimilated with the simplest colloquial language. Has any attempt been made in this direction? Yes, on a small scale in small but possibly growing circles, which have recently begun to publish a periodical in so-called *rōmaji*, i. e. Roman letters. But the dominating tendency of development during the last fifty years has been in the very opposite direction. The acquaintance with western culture, ever since the great reform era, dating from 1868, forced the creation of masses of new words, new terms for western ideas, of a general cultural nature or technical ones. Tens of thousands of new words had to be coined, and here the Japanese let themselves be lured into a false track. They would not borrow the many foreign words, especially Latin and Greek formations, such as 'absolutism', and 'democracy', or 'substantive' and 'phonetic', or 'velocipede' and 'telephone' in forms based on the pronunciation of any western language, but preferred to translate them, just as nowadays Germans are intent upon saying Fernsprecher, Schnellrad, etc. So far, well and good. But they did not translate them by genuine Japanese stems, which would have given somewhat longer and more ungainly, but perfectly clear forms; they created words corresponding to the western ones, with

the help of the short Sino-Japanese words, which here had to play the same part as the Latin and Greek stems had done in the west in the invention of new cultural terms. For the word 'stearin' they coined *kō-shi* 'hard-fat'. But as there are innumerable words *kō*, and innumerable words *shi* among the Sino-Japanese words, we find, in an ordinary lexicon, 27 different words *kō-shi*, one of them meaning 'minister', another 'reflection', another 'white teeth', another 'hind leg', etc. (of course with different Chinese characters for the different words). Owing to this, *kō-shi* for 'stearin' is ambiguous to the ear, and incomprehensible in its purely phonetically written form. If, instead, they had coined a word for 'stearin' from the genuine Japanese stems *katai* 'hard' and *abura* 'fat', they would have got a somewhat longer word, it is true, but one that the ear, and not only the eye, would understand, a word that could therefore have been written phonetically and yet be clear. But, as I have said, into this false track, into this deplorable blind alley, the Japanese have been blindly rushing for the last half-century. The problem of getting rid of the Sino-Japanese words that were ambiguous to the ear was hard enough before 1868. It has now been aggravated ever so much by the myriads of new words of a general cultural or technical nature which have been newly created during these decades with Sino-Japanese word-stems and introduced into Japanese literature. Now that the damage has been done, people in Japan realise that it was a mistake, and desperate attempts are being made to limit the use of the Chinese characters. As thoughtless as they are hopeless have been the attempts made by the authorities to attain this result by reducing the number of Chinese characters taught in the schools. At pre-

sent they are experimenting with a stock of less than 2000 characters. It goes without saying that every sensible schoolboy who, in learning the first thousand characters, becomes quite familiar with the elements of which they are composed, can then, from newspapers, novels, etc., with Kana in the margin to show the pronunciation, easily and without any trouble supply the thousand characters he has not learnt at school. Mechanical restrictions are of no use; something *positive* is wanted here. The next step in the reform — after the first and most important one has been taken, viz. the rejection of Chinese characters for genuine Japanese words, and words that have become fully domesticated — must be a well-directed purist movement. The Sino-Japanese words for general cultural and technical ideas must be replaced by others that the ear can understand, and which can therefore be written phonetically with Kana or letters. The mistakes of the past must be corrected. As far as can conveniently be done, the words should be coined from Japanese word-stems. But there exaggerations must be avoided, or absurdities may result. We may respect the national enthusiasm underlying word creations like 'Fernsprecher' and 'Schnellrad', but, as soon as we are in the domains of technical, grammatical, and scientific phraseology, the words should be as international as possible. I do not know whether the German enthusiasts for purism have tried to create «national» words for 'magnesium oxide' or 'hypercatalectic', or a common word like 'orchid'; but I venture to assert that the Japanese creators of new words would do wisely, in these special domains, to introduce phonetic reproductions of the Latin and Greek words, even if the sound resources do not allow of an exact

resemblance. There, after all, they would be on the basis of a solid, international, clear terminology. The words for more general cultural ideas, which can be more easily made out of purely Japanese word-material, have to be created by literary men; a successful purist movement without capable creative artists at the head, and emanating only from the workshop of the philologist, is to my mind bound to be a failure.

Much of what I have said about linguistic and clerical reforms in Japan, also holds good of China; but there are considerable differences.

\*

If, in Japan, a rational reform of the script means a violent breach with tradition, in China it is tantamount to a radical revolution, for in China two reforms are required, which are independent even if they are logically connected. One is this: away with the old artificial literary language, the Latin of China; replace it with literature based directly on the spoken language. The other is this: away with the old ideography; replace it with phonetic writing. It follows from what I have set forth above in detail, that the latter reform would embrace also the former, for a literary text re-written phonetically would in itself be incomprehensible, owing to the vast number of homonymous words which are only distinguishable by the Chinese script. But the former reform need not perforce include the latter, for there is nothing to prevent present-day colloquial Chinese being written with Chinese characters.

An example. A colloquial sentence like this:

*pa ché-ko shī-fou fang tsai cho-tsī shang*  
(object) this piece stone-head put on table(suffix) top

is quite intelligible to the ear, and can consequently be understood even if written only with Roman letters. But if rendered in the austere literary language:

<i>chī</i>	<i>shī</i>	<i>yū</i>	<i>cho</i>
put	stone	on	table.

it is absolutely incomprehensible by itself, because there are dozens of *chī*, dozens of *shī*, many *yū*, several *cho*. This must perforce be rendered by Chinese characters 置石於棹 in which the 置 *chī* 'put' is clearly differentiated from every other *chī*, etc.

Ergo: if the Chinese are to get rid of the ideographs and use Roman letters, they must also abolish the literary style. But the sacrifice of the literary language does not in itself prevent the preservation of the Chinese script, for the colloquial sentence above, with its 10 syllables, can of course be written with 10 corresponding Chinese characters.

The replacing of the artificial literary language by a new one, based on the modern spoken form of Chinese, is a reform that has already got to work. For a number of years it has had its enthusiastic and courageous champions. They are up against a mighty opposition, and the victory cannot be easy or rapid. The spokesmen of conservatism have weighty arguments to urge, for such a reform would be attended by enormous losses.

In the first place the reform would cut the Chinese off from the national culture, and break the connection with the past. The young Chinese brought up on the new literary language, would lose the faculty of reading the old; and the treasures of thousands of years of literature, in which the whole of Chinese culture from its emergence to

its greatest development, the soul of China, lies hidden, would be to them a closed world. As a substitute they will take to a popular version of western literatures and western culture, without the possibility of really assimilating them, since social conditions in their own country are so entirely different from those of the west.

This argument is sound, but it cannot be decisive. There is, of course a vast array of, on the one hand, sentimental values, love of tradition, national pride, and on the other, practical difficulties, that must be overcome; but if the reformers have sufficient enthusiasm and perseverance, the difficulties can be overcome. Just as surely as the Italians could not go on for ever writing Latin as their literary language, the spoken language bursting forth irresistibly in literature; so here. The study of Chinese antiquity must not cease, but it must assume a different character. The Chinese students must work with their old language as a special branch of study, as the Italians have to work with Latin, and the Norwegians with Old Icelandic. For the man in the street, translations would have to be made into the colloquial language of all the more important documents that contain the quintessence of the Chinese mind<sup>1</sup>. All this will make great claims on the intellectual world of China, but great claims are there in order to be met.

<sup>1</sup>) This is a point of view which, in my opinion, the radical leaders in China pay far too little attention to, and which I would impress upon them. Give the people good translations in Mandarin of the conversations of Confucius, Mencius, certain parts of Liki, Chuangtsi, Tsochuan, Shih-ki, T'ung-kien kang-mu, and a number of other works. Poetry and the essay, literature the value of which lies in its literary style, cannot, of course, be made accessible to the people, but there is a sufficiently great body of literature that can be paraphrased into every-day language without sacrificing much of its value.

Another, and not less important point of view is that breaking the connection with the classical antiquity of China also means breaking the connection between different parts of China. For the artificial literary language is a powerful connecting-link, not only culturally, through the study in different places of the same early literature, but also purely practically, as regards modern writings. The language of literature has become a veritable Chinese esperanto. Differently as people may speak, whatever variants there may be in the pronunciation of words, in the vocabulary, the use of grammatical auxiliaries in the changing dialects; as soon as one expresses oneself in script, the differences disappear, and the pronunciation becomes a matter of indifference: the characters speak directly to the eye, the vocabulary is the same throughout China, carefully confined to a certain selection of words with a literary sanction, which are strictly differentiated from the common, local words; and the grammar is in the main that of Old Chinese. A governmental edict from Peking, a newspaper article from Canton, a textbook compiled in Shanghai are, thanks to this written esperantoism, comprehensible to readers throughout China. For works of this purely practical nature, a simple literary form of language, which has been and is eminently practical, the so-called 'easy *wén-lí*', easy literary style, has actually taken a definite shape. The reformers must send this useful servant packing, and thereby they sever at a stroke the intimate practical connections between the various parts of the country. For a colloquial literature based, for instance, on Pekinese, whether it be written phonetically or with Chinese characters, would be something quite foreign to the Cantonese or Shanghaian. It might be urged that



the Cantonese or the Shanghaian could learn this language just as easily as the artificial literary language; both could be equally foreign to, or equally allied to the Cantonese and Shanghai dialects. This is perfectly true; but these good people will not be willing to do so, for not only does it stop them from killing two birds with one stone, as they have hitherto done in learning at once the language of the old literature *and* a modern esperanto; it also has the disadvantage that the newly created Pekinese literary language cannot have the same weight as the written esperanto, which has a venerable organic unity with the old language. No; the inevitable result would be that every dialect area of any size and importance would create its own modern literary language.

Why, then, do the reformers not favour this written esperanto, simple and normalised, practical and universally comprehensible? In the first place because they want to write as they speak. Then, furthermore, they want the written word to be of such a nature that it can be read aloud and understood, and not be merely a literature for the eye. And finally, experience has taught them that it is a great temptation for many people, when they are writing in the traditional literary language, to use phrases and allusions that enable them to show off their familiarity with the old literature, but which at the same time make their style incomprehensible to anyone but an accomplished scholar. As long as the artificial language is retained, literature will never be truly popular, they say; will never be for the multitude. However this may be, the first two reasons are more than enough. The reformers have a sound fundamental principle in proclaiming that the literature of to-day shall be based on the speech of to-day.

Taking their stand upon that, they can freely declare that those who will not side with their N. Chinese literary creation, may go their own way. Possibly this is only a confirmation in the cultural domain of the fact that grows more and more evident in the political: that the most southerly part of China cannot and will not remain any longer within the frame of Chinese unity. And here it is surely wise, as elsewhere, to look facts in the face, to be in the vanguard of development, and not lag behind and whimper.

Go ahead, then, with a modern literary language based on the colloquial one! But how will it have to look? That is not such an easy question. As a matter of fact, there are several possibilities.

I mentioned earlier that among the numerous dialects in China there is a great group spoken throughout the north and centre, down to the Yang-tsi and somewhat beyond, the Mandarin dialects, which are so like each other that their representatives understand one another fairly well; a group of which the Peking variant is the most fashionable. The reformers take it for granted that it is Mandarin that is to be the new literary language; and a whole series of writings skilfully based on this idiom has already seen the light of day from the Peking professors Hu Shī, Ts'ien Hūan-t'ung, and many others. But hitherto, inasfar as the movement has been seriously literary, it has only dealt with the language itself, its notation being effected by means of the old characters.

By this many advantages have been gained. In the first place, the reformers have not sprung the movement too suddenly on the general public. The shock is severe enough as it is — the shock of rejecting the venerable

literary language, and adopting *pai-hua* 'white language', i. e. the drab, simple, unadorned colloquial language which is regarded as vulgar, not only for ephemeral writing, light reading, as has been done for centuries, but also for serious literature, scientific dissertations, periodicals, and even poetry. Here the preservation of the old script has somewhat softened the brutality in the changes suggested.

Moreover, they have been able to write for, and be immediately understood by, the whole of the North and Centre of China, without being hampered by the subdialects, by the variants in the pronunciation of the signs inside the Mandarin area. The phraseology and vocabulary have, it is true, been of such a nature occasionally as not to be current throughout the area, but this has only affected some isolated points, since the chief dialect differences in this vast territory were but dimly visible behind the wall of the ideographic script.

Finally they have managed by this means to keep the door open, so to speak, towards the old language. For the student who, in connection with the new literary language based on the colloquial, has learnt the necessary three or four thousand characters, it is naturally much easier to start his excursions into the labyrinths of the early literature than for one who only studies a phonetically written modern language.

It would therefore appear that this stage, to which a number of vigorous reformers have already attained — the discarding of the artificial literary style, but the preservation of the old script — were a happy medium in the efforts towards reform. But more radical reformers might raise objections to such half-measures. If the literary language is based on the modern colloquial language, it

should, just as the colloquial language can be grasped by ear, also be writable phonetically and yet be readable. It would then be quite *unnecessary* for the swarming millions of Chinese schoolchildren, who have no intention of all becoming scholars, to cram up thousands of quaint characters, when a few dozen phonetic symbols, letters or other signs could suffice. This is undoubtedly a weighty argument. Another, less obvious but equally important, is this: as long as the Chinese write their literature with the old characters, which carefully distinguish all homophonous words, the many different *i*, the numerous *shī* etc., it is a great temptation for an author, even if he aims at writing a colloquial language, to drop back into a phraseology that appeals to the eye rather than to the ear, with confusing homonyms when read aloud. In other words: Unless there is a complete divorce from the old language, by even discarding the Chinese script and writing phonetically, it will be impossible, in spite of the best intentions, to get rid of the drawbacks that characterise it. Its obscurities so easily creep in, even in the new colloquial literature; not in descriptions of everyday concrete things, but as soon as somewhat more abstract and difficult notions are touched on. We have seen that what characterises the spoken tongue compared with the old literary language, is the fact that, for clarity, a considerable use is made of compound words, where the old language had simple ones. The Chinese are now arriving at the very same situation in which the Japanese have been for the last half-century. Just as they in their foolishness disdained their longer Japanese words, when it was a question of creating new words for new ideas, of a cultural or technical nature, and used in place of them the short Sino-Japanese monosyllabic words, clear to the

eye but ambiguous to the ear; so is the case here. As long as the alluring ideography may be employed, writers will have recourse to the expedient of creating, even in the new literature, terms for the new ideas from the short old simple words; and despite all there will arise new formations incomprehensible to the ear.

Experience already shows that this is a real danger. If we read the writings of the reformers, as they have appeared in various polemic periodicals over there, we shall find, it must be admitted, a spirited and at times quite brilliant diction in modern Pekinese, though written in Chinese characters. But if we transcribe a few passages with ordinary letters, and then try to read them, we are left at a nonplus. Some of it reads well enough, but it fails as soon as we get to abstract and scientific terms, etc., cultural terms in general — they are all made up of a word-material that is too short and concise to be clear to the ear alone, i. e. to be clear in alphabetical script. And in this respect the reform is incontrovertibly a lame half-measure.

The dictum of the radical reformer must be that we shall be enabled to *speak*, not only about commonplace things, but about the highest abstract things, cultural conceptions and technical phenomena, and we must not be seduced by the elucidating ideography to coin terms that cannot be used in *speech*. Let us therefore use a strictly *phonetic* script. If a writer using a phonetic script will create a term that is to be understood by the reader, he is compelled to put it together from available word-material in such a full and clear manner that it can be understood even when it is read aloud. A purely phonetic script is therefore the only guarantee for a literary language that can be fully alive and organically bound up with the colloquial language.

Another, and in my opinion paramount reason for a purely phonetic script, is the ruthless intrusion of western cultural elements and ideas into China. It means that the Chinese must import tens of thousands of new terms for new ideas, partly scientific and technical, partly of a general cultural character. And here the situation is quite on a par with the one we have just described as far as Japan was concerned. To a great extent the new terminology can and should be created by translation. This has largely been done; and as the Japanese were the first in this field, the Chinese have very extensively simply taken over the new formations made by them; the Japanese, as we know, manufactured the new terms with Sino-Japanese words (written in Chinese), and the Chinese have now in their turn only adopted these productions, reading them, of course, in the Chinese way (varying according to dialect). But as I pointed out earlier, there is a limit to translation, which should not be exceeded unless one is anxious to become ridiculous and unpractical. When in desperation the Chinese use translations for 'dichotyledons' or 'homousians', it is because the Chinese script is so ill adapted to reproducing foreign combinations of sounds, and they have therefore avoided transcriptions as much as possible. In a new, phonetically written literary language it would be easy and elegant, in dealing with such technical terms as ought to be international, to insert them written in a neutral Latin (or Greek) version — or, why not in the form these international words have in Esperanto or Ido?

This is true, to an even greater extent, of the myriads of foreign names that, in spite of all, the Chinese now have to deal with — these cannot be translated! Now when these are rendered by means of clumsy combinations of

Chinese characters, when Verdun becomes *Fan-ör-tun*, Alcibiades *Ya-chi-pi-té*, Krupp, *K'o-lu-po*, it is the merest chance if one succeeds in guessing what name is aimed at; through this difficulty, newspaper articles, especially about foreign things, become preposterous. This difficulty disappears at once with the introduction of phonetic script, in which foreign names can be represented in an unmutilated condition.

There is a middle way between the old script and the purely phonetic, which I will mention en passant, and that is, writing the old Chinese characters as hitherto, but adding to them in their right margin their pronunciation with tiny sound symbols, as has been done in Japanese (Kana), a system for which considerable propaganda has been made the last few years. It is clear that this does not solve the difficulties. In Japanese, where every character can be read in several, sometimes in many ways, it is sorely needed. In Chinese, where a character as a rule only has *one* reading in Pekinese, there is very little object in it. *Either* the language is such that it can be written phonetically and yet become unambiguous and readable; in which case it is best to write purely phonetically, and leave the old texts in Chinese characters to the learned; *or* the language is such that it becomes ambiguous when written phonetically (owing to the homophones), and then we must retain the old characters that distinguish the homophones. In the latter case we get no help, if we do not know a certain character, and do not know its meaning, by being told in the side signs that it is to be read, let us say, *chu*; we do not know which of the innumerable *chu* is meant. If we do know the word, on the other hand, with character, meaning and all, then the side sign is unnecessary. The system has

therefore, as far as I can see, a future exclusively as an aid in primary school teaching, where it could be of great help.

The champion for a new literary language but the old script, however, will raise more objections. If he is to write in a purely phonetic way, he must choose some definite Mandarin dialect, and then he will not be understood by the whole Mandarin area, as he would be if he wrote ideographs. The different pronunciations of dialect speakers would make it hard for them to understand a phonetic text in a Mandarin subdialect (e. g. Pekinese) other than their own (e. g. Nankinese).

To this the radical reformer should answer: «No, it is not so bad as all that. For if you write in Pekinese, and are to be understood by a Central Chinaman, he can learn the *equivalents* in a few hours, the system of phonetic correspondence, e. g. that you always write *chi*, *chiang*, *chiao* in the words he pronounces *tsi*, *tsiang*, *tsiao*, etc.; whereas it would take him a few years to learn the Chinese characters. Moreover, there is every reason to hope that if a great literature arises in phonetically written Pekinese, which would be read throughout the Mandarin area, it would gradually influence the educated spoken language throughout that area; the differences between the various Mandarin dialects would be levelled down, and we should get what never existed before, viz. a single educated colloquial language for two-thirds of that enormous country.»

This must be taken with a slight reserve. In the essential features the basis should, of course, be a definite, concrete, living dialect, therefore most naturally Pekinese; but one should not on that account go too far, and introduce into the literature extreme Pekinisms, Pekinese vulgarisms in vocabulary and pronunciation, which would sound strange



or objectionable as soon as they got beyond the environs of the capital. Normalising to some extent, in order to agree as far as possible with a considerable number of Mandarin variants, is not to be despised, but rather to be recommended. The creation of a fairly neutral written medium, which has *not too much* local colour, is one of the most important tasks for a civilised nation. It must be closely enough allied to a living language to be a really organic, living language itself, but it must throw a bridge across to closely allied dialects, by wise advances in the matter of spelling and a well thought out moulding of the vocabulary, keeping as far as possible to words that are widely and generally current.

The radical reformers in China have heeded this claim for adjustment to Mandarin in its entirety, and here they have allowed themselves, in my opinion, to be enticed too far. They have set up an artificial Mandarin language, a kind of mean proportional between the different dialects, to which they have given the name of *kuo-yü* 'the national language', and have tried to launch this not only as the written but also as the spoken language. Such an attempt is, I am convinced, foredoomed to failure. I do not believe that an artificial language, made by theorists, can ever become the natural language of a large group of people. You can force upon a group of people a language different from their original one, but the new language must then be one which is in use, and which is the mother tongue of another group of people — not an artificial language arrived at by a theoretical procedure and by generalising. The salient point in the reform movement is consequently this: will the young reformers manage to create a written medium which, on the one hand, is so intimately based

upon Pekinese that it will be a living child and not an artificial doll, and which, on the other hand, by paying some regard to the historical factor (i. e. by a spelling that meets all the Mandarin dialects to some extent, and a strict bridle on the vocabulary), promises to be a real literary language? Undoubtedly they will; it is only a question of time.

Purely practical difficulties of various kinds will, moreover, crop up. If we are to write this modified Peking-Mandarin with letters like a western language, will it be possible to reproduce the Chinese sounds exactly with our letters? Of course not. For this purpose one should, properly speaking, in addition to the ordinary letters, adopt a number of symbols taken from phonetic alphabets. But one must not be too scientific: this would create difficulties, especially of a typographical nature. To begin with, one must therefore tamper somewhat with accuracy, and proclaim once for all what values are to be assigned to the western letters in the written Chinese, and then not worry if they deviate a little from western use. When *ken* is written, it should have to be taken for granted that the *e* here has not the ordinary *e*- or *ä*-sound (as in German *mehr*, *Bär*), but has to be pronounced very much like the vocal murmur in French *le* (*lə*). When *hi* is written, it should be understood that the *h*- is pronounced here like a German *ich*-laut (almost as in English *huge*).

There is no harm in this (even if a few special phonetic symbols were the ideal). So far, so good. But there are more serious difficulties. Our set of letters are fairly good for our western languages, which as a rule consist of consonants and vowels only. But, as mentioned above, Chinese words have a third element, which is just as important, and just as necessary an ingredient in a word as its vowels

and consonants; and that is the tones, the musical accent. A text written with letters which does not distinguish the tones of the words, which writes *i* for *ī*, *i'*, *i<sup>~</sup>* and *i'*, is unreadable. The tones must therefore be marked. But how? There are many systems. The commonest are to put accents over the vowels, or to put figures in the right hand upper corners of the syllables. But can we reasonably expect that the Chinese, in order to have the honour of adopting our western alphabetic writing, which consequently is only partially adaptable to Chinese, shall put up with writing diacritic signs on every syllable? Are they to be compelled to write sentences which look like this:

*fā nà-kò p'eng-yú, wó pù hên hī-huān,*

or: *f<sup>a1</sup> na<sup>4</sup>-ko<sup>4</sup> p'eng<sup>2</sup>-yu<sup>3</sup>, wo<sup>3</sup> pu<sup>4</sup> hen<sup>3</sup> hi<sup>3</sup>-huan<sup>1</sup>.*

This is clearly unreasonable. The young phoneticians out there are consequently experimenting with various devices to enable them to write with unmodified letters and yet with the tones marked — the programme is: Pekinese shall be written with tones and all, and the symbols shall be those that exist on an American type-writer and no more. For this they are working with various spellings of one and the same syllable according to the tones it has; for instance:

<i>iaū</i>	is written <i>iau</i> ,	<i>t'anḡ</i>	is written <i>t'ang</i>
<i>iau'</i>	» <i>yau</i>	<i>t'ang'</i>	» <i>t'arng</i>
<i>iau<sup>~</sup></i>	» <i>eau</i>	<i>t'ang<sup>~</sup></i>	» <i>t'aang</i>
<i>iau'</i>	» <i>iaw</i>	<i>t'ang'</i>	» <i>t'anng</i>

and so on. The obvious danger is that they get so far away from the phonetic truth, and create so artificial a system that it cannot succeed. The last word is far from being spoken in the matter, and many more experiments are required.

It lies outside the scope of this sketch to deal in detail with all the extremely complicated problems in the reforms of language and script in the far East. It has only been my intention to give the reader a slight idea of what a dreadful skein of problems it is. The leading men will require much discretion and suffer many bitter experiences, before they can steer successfully through the skerries. But, above all, it needs something I have scarcely touched on. It needs *great men*. Language, literary language, is the greatest artistic creation of a civilized nation; it is not shaped by philologists at their desks, but by the giants of thought who have something to say to their contemporaries, and who carve their monuments from the solid granite of the spoken language.

But for this very reason I look with confidence to the future of Chinese literature. China has always had artists, great artists, in every sphere; and when the throes of these critical years have been lived through, Chinese literature will arise like a phenix from its ashes, in new strength and beauty. Of this no one can doubt who has learnt to know, to understand, and to admire the Chinaman in the past and in the present.

---